

Initial Score = 175 Optimized Score = 175 Significance = 0.57
Residue Identity = 84% Matches = 175 Mismatches = 33
Gaps = 0 Conservative Substitutions = 0

```
X
TCAGAGAACGCGCTACGACCTGGCCGGGTGGCCGCTACAGGTCACAAAGAACTGGGCTGAACACCG
|||||
TCAGAGAACGCGCTACGACCTGGCCGGGTGGCCGCTACAGGTCACAAAGAACTGGGCTGGACGTCG
X
10 20 30 40 50 60 70
```

```
80 90 100 110 120 130 140
AGAAATGCCCCAACCCACGACGACCTGACCCGAGAGAGACGTCGTGCTCCACCATGCAATACCTGGTCCCTTGC
|||||
GCGATCCGATACACCACTCCACGCTGACCGAGAGACGTCGTGCTCCACCATGCAATACCTGGTCCCTTGC
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200
ACGAGGCGACGCGCCACGATGAAAGTCCCGGTGGCGCTGACAGTCCGCTGAGACCGACACAT
|||||
ACGAGGCGACGCGCCACGATGAAAGTCCCGGTGGCGCTGACAGTCCGCTGAGACCGACACAT
150 160 170 180 190 200 X
```

9. US-09-697-123B-24 (1-208)
US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 174 Optimized Score = 174 Significance = 0.55
Residue Identity = 83% Matches = 174 Mismatches = 34
Gaps = 0 Conservative Substitutions = 0

```
X
TCAGAGAACGCGCTACGACCTGGCCGGGTGGCCGCTACAGGTCACAAAGAACTGGGCTGAACACCG
|||||
TCAGAGAACGCGCTACGACCTGGCCGGGTGGCCGCTACAGGTCACAAAGAACTGGGCTGAACACCG
X
10 20 30 40 50 60 70
```

```
80 90 100 110 120 130 140
AGAAATGCGCCAAACCCACGACGACCTGACCCGAGAGAGACGTCGTGCTCCACCATGCAATACCTGGTCCCTTGC
|||||
GCCACGCCCATCACCAGCTGACCGCTGACCCGAGAGAACGTCGTGCTCCACCATGCAATACCTGGTCCCTTGC
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200
ACGAGGCGACGCGCCACGATGAAAGTCCCGGTGGCGCTGACAGTCCGCTGAGACCGACACAT
|||||
ACGAGGCGACGCGCCACGATGAAAGTCCCGGTGGCGCTGACAGTCCGCTGAGACCGACACAT
150 160 170 180 190 200 X
```

10. US-09-697-123B-24 (1-208)
US-09-697-123B-2 Sequence 23, Application US/09697123B

Initial Score = 173 Optimized Score = 173 Significance = 0.53
Residue Identity = 83% Matches = 173 Mismatches = 35
Gaps = 0 Conservative Substitutions = 0

```
X
TCAGAGAACGCGCTACGACCTGGCCGGGTGGCCGCTACAGGTCACAAAGAACTGGGCTGAACACCG
|||||
TCAGAGAACGCGCTACGACCTGGCCGGGTGGCTGCTTACAGGTCACAAAGAACTGGGCTGGACGCGC
X
10 20 30 40 50 60 70
```

```
80 90 100 110 120 130 140
AGAAATGCGCCAAACCCACGACGACCTGACCCGAGAGAGACGTCGTGCTCCACCATGCAATACCTGGTCCCTTGC
|||||
GTGAGCGCGATCAAGAGTCCGACGTCGAGAGAGACGTCGTGCTCCACCATGCAATACCTGGTCCGCTTGC
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200
ACGAGGCGACGCGCCACGATGAAAGTCCCGGTGGCGCTGACAGTCCGCTGAGACCGACACAT
|||||
ATGAGGCGTACGTGACGATGACCGTTCACAGGTGGGTGAGAGTCCGATGATACCTGTCGACGACAT
150 160 170 180 190 200 X
```

11. US-09-697-123B-24 (1-208)
US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 173 Optimized Score = 173 Significance = 0.53
Residue Identity = 83% Matches = 173 Mismatches = 35
Gaps = 0 Conservative Substitutions = 0

```
X
TCAGAGAACGCGCTACGACCTGGCCGGGTGGCCGCTACAGGTCACAAAGAACTGGGCTGAACACCG
|||||
TCAGAGAACGCGCTACGACCTGGCCGGGTGGCCGCTACAGGTCACAAAGAACTGGGCTGGACGTCG
X
10 20 30 40 50 60 70
```

```
80 90 100 110 120 130 140
AGAAATGCGCCAAACCCACGACGACCTGACCCGAGAGAGACGTCGTGCTCCACCATGCAATACCTGGTCCCTTGC
|||||
GCGATCCGATACACCACTCCACGCTGACCGAGAGACGTCGTGCTCCACCATGCAATACCTGGTCCCTTGC
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200
ACGAGGCGACGCGCCACGATGAAAGTCCCGGTGGCGCTGACAGTCCGCTGAGACCGACACAT
|||||
ACGAGGCGACGCGCCACGATGAAAGTCCCGGTGGCGCTGACAGTCCGCTGAGACCGACACAT
150 160 170 180 190 200 X
```

12. US-09-697-123B-24 (1-208)
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 172 Optimized Score = 172 Significance = 0.51
Residue Identity = 82% Matches = 172 Mismatches = 36
Gaps = 0 Conservative Substitutions = 0

```
X
TCAGAGAACGCGCTACGACCTGGCCGGGTGGCCGCTACAGGTCACAAAGAACTGGGCTGAACACCG
|||||
TCAGAGAACGCGCTACGACCTGGCCGGGTGGCGCTGCTTAAAGTCAAAAGAACTGGGCTGCAATGTCG
X
10 20 30 40 50 60 70
```

```
80 90 100 110 120 130 140
AGAAATGCGCCAAACCCACGACGACCTGACCCGAGAGAGACGTCGTGCTCCACCATGCAATACCTGGTCCCTTGC
|||||
GCCACGCCCATCACCAGCTGACCGCTGACCCGAGAGAACGTCGTGCTCCACCATGCAATACCTGGTCCCTTGC
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200
ACGAGGCGACGCGCCACGATGAAAGTCCCGGTGGCGCTGACAGTCCGCTGAGACCGACACAT
|||||
ACGAGGCGTACAGCACGATGATGTCGTCGGGCGCGCTGAGAGTCCCGGTGGAACCGACACAT
150 160 170 180 190 200 X
```

13. US-09-697-123B-24 (1-208)
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 172 Optimized Score = 172 Significance = 0.51
Residue Identity = 82% Matches = 172 Mismatches = 36
Gaps = 0 Conservative Substitutions = 0

```
X
TCAGAGAACGCGCTACGACCTGGCCGGGTGGCCGCTACAGGTCACAAAGAACTGGGCTGAACACCG
|||||
TCAGAGAACGCGCTACGACCTGGCCGGGTGGCTGCTTAAAGTCAAAAGAACTGGGCTGCAATGTCG
X
10 20 30 40 50 60 70
```

```
80 90 100 110 120 130 140
AGAAATGCGCCAAACCCACGACGACCTGACCCGAGAGAGACGTCGTGCTCCACCATGCAATACCTGGTCCCTTGC
|||||
GCGAGCCCATACCTGCTGACGCTGACCGAGAGAGACGTCGTGCTCCACCATGCAATACCTGGTCCGCTTGC
80 90 100 110 120 130 140
```

19. US-09-697-123B-24 (1-208)
US-09-697-123B-4 Sequence 4, Application US/09697123B

Initial Score = 138 Optimized Score = 167 Significance = -0.21
Residue Identity = 82% Matches = 172 Mismatches = 35
Gaps = 1 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAAACCGCTACGACCTGGCCGGGTCGCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
|||||
TCAAGAGAAACCGCTACGACCTGGCCGGGTCGCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
X      10      20      30      40      50      60      70
```

```
AGATTCGCCACCAACCAACCACTTGCACCAAGAGAGAGCTGCTGCCACCATCGAATACCTGTGCGCTTGC
|||||
GCGATTCGCATTCACGACCTGCACCGCTGACCAAGAGAGAGCTGCTGCCACCATCGAATACCTGTGCGCTTGC
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200
ACGAGGGGACCGCCACCATGAAGTCCCGCGGTGGCGCTGACAGTGCCTGGTGAAGACCGACGACAT
|||||
ACGAGGGGACCGCCACCATGAAGTCCCGCGGTGGCGCTGACAGTGCCTGGTGAAGACCGACGACAT
150      160      170      180      190      200
```

20. US-09-697-123B-24 (1-208)

US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 136 Optimized Score = 160 Significance = -0.25
Residue Identity = 76% Matches = 170 Mismatches = 38
Gaps = 15 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAAACCGCTACGACCTGGCCGGGTCGCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
|||||
TCAAGAGAAACCGCTACGACCTGGCCGGGTCGCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
X      10      20      30      40      50      60      70
```

```
AGATTCGCCACCAACCAACCTTGCACCAAGAGAGAGCTGCTGCCACCATCGAATACCTGTGCGCTTGC
|||||
GCGATTCGCATTCACGACCTGCACCGCTGACCAAGAGAGAGCTGCTGCCACCATCGAATACCTGTGCGCTTGC
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200
ACGAGG-----GGCAGCGCCACAGTGAAGTCCCGGTGGCGTGAAGTGCCTGGTGAACACCG
|||||
ACGAGG-----GGCAGCGCCACAGTGAAGTCCCGGTGGCGTGAAGTGCCTGGTGAACACCG
150      160      170      180      190      200
```

```
X
ACGACAT
ACGACAT
220 X
```

21. US-09-697-123B-24 (1-208)

US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score = 123 Optimized Score = 161 Significance = -0.53
Residue Identity = 78% Matches = 168 Mismatches = 40
Gaps = 6 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAAACCGCTACGACCTGGCCGGGTCGCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
|||||
TCAAGAGAAACCGCTACGACCTGGCCGGGTCGCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
X      10      20      30      40      50      60      70
```

```
AGATTCGCC-----AACCAACCAACCACTTGCACCAAGAGAGAGCTGCTGCCACCATCGAATACCTGTG
|||||
CCAATTCGCGCTACGACCTGCACCGCTGACCAAGAGAGAGCTGCTGCCACCATCGAATACCTGTGCTGC
80      90      100      110      120      130      140
```

```
140      150      160      170      180      190      200
GCTTGCACGAGAGGGGACCGCCACCATGAAGTCCCGGTGGCGCTTACAAAGTTAAACAAGAACTCGGGCTGA
|||||
GCTTGCACGAGAGGGGACCGCCACCATGAAGTCCCGGTGGCGCTTACAAAGTTAAACAAGAACTCGGGCTGA
150      160      170      180      190      200      210
```

22. US-09-697-123B-24 (1-208)

US-09-697-123B-1 Sequence 18, Application US/09697123B

Initial Score = 119 Optimized Score = 161 Significance = -0.61
Residue Identity = 79% Matches = 167 Mismatches = 41
Gaps = 3 Conservative Substitutions = 0

```
X      10      20      30      40      50      60
TCAAGAGAAACCGCTACGACCTGGCCGGGTCGCTTACAAAGTTAAACAAGAACTCGGGCTGA-----ACA
|||||
TCAAGAGAAACCGCTACGACCTGGCCGGGTCGCTTACAAAGTTAAACAAGAACTCGGGCTGA-----ACA
X      10      20      30      40      50      60
```

```
70      80      90      100      110      120      130      140
CCGAGATTCGCCACCAACCAACCACTTGCACCAAGAGAGAGCTGCTGCCACCATCGAATACCTGTGCGCT
|||||
ACCGAGATTCGCCACCAACCAACCACTTGCACCAAGAGAGAGCTGCTGCCACCATCGAATACCTGTGCGCT
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200
TTCACGAGAGGGGACCGCCACCATGAAGTCCCGGTGGCGCTGACAGTGCCTGGTGAAGACCGACGACAT
|||||
TTCACGAGAGGGGACCGCCACCATGAAGTCCCGGTGGCGCTGACAGTGCCTGGTGAAGACCGACGACAT
150      160      170      180      190      200
```

23. US-09-697-123B-24 (1-208)

US-09-697-123B-2 Sequence 21, Application US/09697123B

Initial Score = 111 Optimized Score = 166 Significance = -0.78
Residue Identity = 80% Matches = 173 Mismatches = 35
Gaps = 6 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAAACCGCTACGACCTGGCCGGGTCGCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
|||||
TCAAGAGAAACCGCTACGACCTGGCCGGGTCGCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
X      10      20      30      40      50      60      70
```

```
AGATTCGCC-----AACCAACCAACCACTTGCACCAAGAGAGAGCTGCTGCCACCATCGAATACCTGTGTC
|||||
CCAATTCGCGCTACGACCTGCACCGCTGACCAAGAGAGAGCTGCTGCCACCATCGAATACCTGTGCTGC
80      90      100      110      120      130      140
```

```
140      150      160      170      180      190      200
GCTTGCACGAGAGGGGACCGCCACCATGAAGTCCCGGTGGCGCTGACAGTGCCTGGTGAACACCGACAT
|||||
GCTTGCACGAGAGGGGACCGCCACCATGAAGTCCCGGTGGCGCTGACAGTGCCTGGTGAACACCGACAT
150      160      170      180      190      200      210
```

24. US-09-697-123B-24 (1-208)

US-09-697-123B-1 Sequence 14, Application US/09697123B

Initial Score = 109 Optimized Score = 171 Significance = -0.82
Residue Identity = 83% Matches = 178 Mismatches = 30
Gaps = 6 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAAACCGCTACGACCTGGCCGGGTCGCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
|||||
TCAAGAGAAACCGCTACGACCTGGCCGGGTCGCTTACAAAGTTAAACAAGAACTCGGGCTGAACACCG
X      10      20      30      40      50      60      70
```

```
      80      90      100      110      120      130
AGATTGCG-----CCAACCCAGCAGCCCTGACCGAAGAGGAGCGTGGCCGACCATCGAATACCTGTGC
      |      |      |      |      |      |      |      |      |      |
CCGACTCGGCGCTACCCGCTCTGACACCGCTGACGAGGATGTGTCGCCACCATCGAGTACCTGTGC
      80      90      100      110      120      130      140
140      150      160      170      180      190      200
GCTTGCAGCAGCGGCGACGCGCATGATGAAGTCCCGCGTGGCGCTGAGAGGTGCCGCTGAGACCGAGACAT
      |      |      |      |      |      |      |      |      |      |
GCCTGCAGCAGCGGCGCATGATGACGCTTCCCGCGCGCTGAGAGGTGCCGCTGAGACCGAGACAT
150      160      170      180      190      200      210
X
TCAGGAGAGAAGCGCTACGA
|||||
TCAGGAGAGAAGCGCTACGA
X
10
X
```

25. US-09-697-123B-24 (1-208)
US-09-697-123B-2 Sequence 25, Application US/09697123B

Initial Score	-	19	Optimized Score	-	19	Significance	-	-2.72
Residue Identity	"	100%	Matches	"	19	Mismatches	"	0
Gaps	"	0	Conservative Substitutions	"			"	0

X
TCAGGAGAGAAGCGCTACGA
|||||
TCAGGAGAGAAGCGCTACGA
X
10
X

> 0 <
0110 IntelliGenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-697-123b-23.res made by shanley on Wed 13 Nov 102 14:13:07 PST.

Query sequence being compared: US-09-697-123B-23 (1-208)
Number of sequences searched: 26
Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123B-23 (1-208) with:
File: US09697123B.seq

100-
N -
U 50-
M -
B -
E -
R -
O -
F 10-
S -
E 5-
Q -
U -
N -
C -
E -
S -
SCORE 0 23 46 69 92 116 139 162 185 208
STDDEV -2 -1 0 1

PARAMETERS
Similarity matrix Unitary 1 K-tuple 4
Mismatch penalty 5.00 Joining penalty 30
Gap penalty 0.33 Window size 205
Gap size penalty 12
Cutoff score 0
Randomization group 0

SEARCH STATISTICS

Scores: Mean 152 Median 177 Standard Deviation 48.36
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 5077
Number of sequences searched: 26
Number of scores above cutoff: 25

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.	Length	Score	Sig.	Frame
---------------	-------------	------------	--------	-------	------	-------

1. US-09-697-123B-2 Sequence 23, Application 208 208 208 1.16 0
The list of other best scores is:

Sequence Name	Description	Length	Score	Init. Opt.	Sig.	Frame
2. US-09-697-123B-1 Sequence 15, Application 208		188	188	0.74	0	
3. US-09-697-123B-2 Sequence 2, Application U 208		185	185	0.68	0	
4. US-09-697-123B-3 Sequence 3, Application U 208		183	183	0.64	0	
5. US-09-697-123B-1 Sequence 16, Application U 208		182	182	0.62	0	
6. US-09-697-123B-6 Sequence 6, Application U 208		182	182	0.62	0	
7. US-09-697-123B-5 Sequence 5, Application U 208		182	182	0.62	0	
8. US-09-697-123B-2 Sequence 22, Application U 208		181	181	0.60	0	
9. US-09-697-123B-1 Sequence 1, Application U 208		181	181	0.60	0	
10. US-09-697-123B-1 Sequence 11, Application U 208		180	180	0.58	0	
11. US-09-697-123B-1 Sequence 10, Application U 208		179	179	0.56	0	
12. US-09-697-123B-8 Sequence 8, Application U 208		177	177	0.52	0	
13. US-09-697-123B-1 Sequence 19, Application U 208		176	176	0.50	0	
14. US-09-697-123B-1 Sequence 17, Application U 208		176	176	0.50	0	
15. US-09-697-123B-2 Sequence 12, Application U 208		173	173	0.43	0	
16. US-09-697-123B-1 Sequence 24, Application U 207		152	177	0.00	0	
17. US-09-697-123B-4 Sequence 4, Application U 207		148	175	-0.08	0	
18. US-09-697-123B-2 Sequence 20, Application U 205		147	171	-0.10	0	
19. US-09-697-123B-9 Sequence 9, Application U 223		138	163	-0.29	0	
20. US-09-697-123B-1 Sequence 13, Application U 223		136	159	-0.33	0	
21. US-09-697-123B-1 Sequence 14, Application U 214		132	169	-0.41	0	
22. US-09-697-123B-2 Sequence 21, Application U 214		117	161	-0.72	0	
23. US-09-697-123B-7 Sequence 7, Application U 214		117	155	-0.72	0	
24. US-09-697-123B-1 Sequence 18, Application U 211		113	165	-0.81	0	
25. US-09-697-123B-2 Sequence 25, Application U 19		19	19	-2.75	0	

1. US-09-697-123B-23 (1-208)
US-09-697-123B-2 Sequence 23, Application US/09697123B

Initial Score = 208 Optimized Score = 208 Significance = 1.16
Residue Identity = 100% Matches = 208 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCGGGTGTCTTCAAGTCAAGAAAGCTGGTTGCACGGC
|||||
TCAAGAGAAAGCGCTACGACCTGGCGGGTGTCTTCAAGTCAAGAAAGCTGGTTGCACGGC
X 10 20 30 40 50 60 70
GTGAGCGGATCAGCAGCTCGACGCTGACCGGAAAGAGAGAGCGTCCACCATGAGTACCTGTCGGCTGC
|||||
GTGAGCGGATCAGCAGCTCGACGCTGACCGGAAAGAGAGAGCGTCCACCATGAGTACCTGTCGGCTGC
80 90 100 110 120 130 140
ATGAGGTCACGTGACGATGACGCTTCCAGTGGCCGTCGAGTGGAGTGGATGATGACACAT
|||||
ATGAGGTCACGTGACGATGACGCTTCCAGTGGCCGTCGAGTGGAGTGGATGATGACACAT
150 160 170 180 190 200
X

2. US-09-697-123B-23 (1-208)
US-09-697-123B-1 Sequence 15, Application US/09697123B

Initial Score = 188 Optimized Score = 188 Significance = 0.74
Residue Identity = 90% Matches = 188 Mismatches = 20
Gaps = 0 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCGGGTGTCTTCAAGTCAAGAAAGCTGGTTGCACGGC
|||||
TCAAGAGAAAGCGCTACGACCTGGCGGGTGTCTTCAAGTCAAGAAAGCTGGTTGCACGGC
X 10 20 30 40 50 60 70

8. US-09-697-123B-23 (1-208)
US-09-697-123B-2 Sequence 22, Application US/09697123B

80 90 100 110 120 130 140

X	10	20	30	40	50	60	70
TCAGGAGAGCGCTACGACCTGGCCCGGCTGGTCTTACAAAGTCACAGAAGCTTGGGCTTGACAGCCG							

> 0 <
0110 Intelligenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-697-123b-22.res made by shanley on Wed 13 Nov 102 14:12:29 PST.

Query sequence being compared: US-09-697-123B-22 (1-208)
Number of sequences searched: 26
Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123B-22 (1-208) with:
File: US09697123B.seq

```
100- - - - -
N - - - - -
U - 50- - - -
M - - - - -
B - - - - -
E - - - - -
R - - - - -
O - 10- - - -
F - - - - -
S - - - - -
E - 5- - - -
O - - - - -
U - - - - -
N - - - - -
C - - - - -
S - - - - -
SCORE 0 231 46 69 92 116 139 162 185 208
STDEV -2 -1 -1 0 1 1 1 1 1 1
```

PARAMETERS

Similarity matrix Unitary 1 K-tuple 4
Mismatch penalty 1 Joining penalty 30
Gap penalty 5.00 Window size 205
Gap size penalty 0.33
Cutoff score 12
Randomization group 0

SEARCH STATISTICS

Scores: Mean 155 Median 178 Standard Deviation 51.71
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 5077

Number of sequences searched: 26

Number of scores above cutoff: 25

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.	Length	Score	Sig.	Frame
---------------	-------------	------------	--------	-------	------	-------

1. US-09-697-123B-2 Sequence 22, Application 208 208 208 1.02 0
The list of other best scores is:

Sequence Name	Description	Length	Score	Opt.	Sig.	Frame
2. US-09-697-123B-1 Sequence 16, Application 208	207	207	1.01	0		
3. US-09-697-123B-6 Sequence 6, Application 208	207	207	1.01	0		
4. US-09-697-123B-5 Sequence 5, Application 208	207	207	1.01	0		
5. US-09-697-123B-1 Sequence 10, Application 208	185	185	0.58	0		
6. US-09-697-123B-1 Sequence 11, Application 208	184	184	0.56	0		
7. US-09-697-123B-3 Sequence 11, Application 208	184	184	0.56	0		
8. US-09-697-123B-2 Sequence 15, Application 208	183	183	0.54	0		
9. US-09-697-123B-2 Sequence 23, Application 208	181	181	0.50	0		
10. US-09-697-123B-1 Sequence 19, Application 208	181	181	0.50	0		
11. US-09-697-123B-8 Sequence 8, Application 208	181	181	0.50	0		
12. US-09-697-123B-1 Sequence 1, Application 208	181	181	0.50	0		
13. US-09-697-123B-2 Sequence 2, Application 208	180	180	0.48	0		
14. US-09-697-123B-1 Sequence 17, Application 208	177	177	0.43	0		
15. US-09-697-123B-2 Sequence 24, Application 208	172	172	0.33	0		
16. US-09-697-123B-2 Sequence 20, Application 205	160	170	0.10	0		
17. US-09-697-123B-1 Sequence 12, Application 207	156	175	0.02	0		
18. US-09-697-123B-4 Sequence 4, Application 207	153	178	-0.04	0		
19. US-09-697-123B-9 Sequence 9, Application 223	143	159	-0.23	0		
20. US-09-697-123B-1 Sequence 13, Application 223	136	161	-0.37	0		
21. US-09-697-123B-7 Sequence 7, Application 214	120	158	-0.68	0		
22. US-09-697-123B-1 Sequence 18, Application 211	114	158	-0.79	0		
23. US-09-697-123B-1 Sequence 14, Application 214	109	167	-0.89	0		
24. US-09-697-123B-2 Sequence 21, Application 214	108	161	-0.91	0		
25. US-09-697-123B-2 Sequence 25, Application 19	19	19	-2.63	0		

1. US-09-697-123B-22 (1-208)
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 208 Optimized Score = 208 Significance = 1.02
Residue Identity = 100% Matches = 208 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

TCAGAGAGAAAGCGCTACGACCTGCGCCGCGTGGCTATTAAGTCAACAAAGAGTGGCGCATGTCG
TCAGAGAGAAAGCGCTACGACCTGCGCCGCGTGGCTATTAAGTCAACAAAGAGTGGCGCATGTCG
X 10 20 30 40 50 60 70

GGGAGCCCATCTACGTCGTGAGCGCTGACCGAAGAGAGAGCGTGGCCACCATTCGAATATCTGGTCGCTGC
GGGAGCCCATCTACGTCGTGAGCGCTGACCGAAGAGAGAGCGTGGCCACCATTCGAATATCTGGTCGCTGC
X 80 90 100 110 120 130 140

ACGAGGCTCAGACCGATGATGCTTCGCGCGCGTGGAGCTCCGGTGAACCCGACGACAT
ACGAGGCTCAGACCGATGATGCTTCGCGCGCGTGGAGCTCCGGTGAACCCGACGACAT
X 150 160 170 180 190 200

2. US-09-697-123B-22 (1-208)
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 207 Optimized Score = 207 Significance = 1.01
Residue Identity = 99% Matches = 207 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

TCAGAGAGAAAGCGCTACGACCTGCGCCGCGTGGCTATTAAGTCAACAAAGAGTGGCGCATGTCG
TCAGAGAGAAAGCGCTACGACCTGCGCCGCGTGGCTATTAAGTCAACAAAGAGTGGCGCATGTCG
X 10 20 30 40 50 60 70

US-09-697-123B-1 Sequence 15, Application US/09697123B

Initial Score = 183 Optimized Score = 183 Significance = 0.54
Residue Identity = 87% Matches = 183 Mismatches = 25
Gaps = 0 Conservative Substitutions = 0

X
TCGAAGGAGAAAGGCTACGACCTGGCCCGGCGTGGTGGTATTAAGTCAACAAAGAGCTCGGCTGCATGTGC
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
10 20 30 40 50 60 70
100 110 120 130 140

CGAGCGCCATCAGCTGCTGACGCTGACCGGAGAGAGAGCTGCGCCACCATCGAATATCTGCTCCCTTGC
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
80 90 100 110 120 130 140
150 160 170 180 190 200

ACGAGGCTCAGACACGAGATGATCTCCGGCGGCGTGGAGTGGGAGAAACCGACGACAT
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
150 160 170 180 190 200

9. US-09-697-123B-22 (1-208)

US-09-697-123B-2 Sequence 23, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.50
Residue Identity = 87% Matches = 181 Mismatches = 27
Gaps = 0 Conservative Substitutions = 0

X
TCGAAGGAGAAAGGCTACGACCTGGCCCGGCGTGGTGGTATTAAGTCAACAAAGAGCTCGGCTGCATGTGC
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
10 20 30 40 50 60 70
100 110 120 130 140

CGAGCGCCATCAGCTGCTGACGCTGACCGGAGAGAGAGCTGCGCCACCATCGAATATCTGCTCCCTTGC
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
80 90 100 110 120 130 140
150 160 170 180 190 200

ACGAGGCTCAGACACGAGATGATCTCCGGCGGCGTGGAGTGGGAGAAACCGACGACAT
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
150 160 170 180 190 200

10. US-09-697-123B-22 (1-208)

US-09-697-123B-1 Sequence 19, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.50
Residue Identity = 87% Matches = 181 Mismatches = 27
Gaps = 0 Conservative Substitutions = 0

X
TCGAAGGAGAAAGGCTACGACCTGGCCCGGCGTGGTGGTATTAAGTCAACAAAGAGCTCGGCTGCATGTGC
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
10 20 30 40 50 60 70
100 110 120 130 140

CGAGCGCCATCAGCTGCTGACGCTGACCGGAGAGAGAGCTGCGCCACCATCGAATATCTGCTCCCTTGC
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
80 90 100 110 120 130 140
150 160 170 180 190 200

ACGAGGCTCAGACACGAGATGATCTCCGGCGGCGTGGAGTGGGAGAAACCGACGACAT
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
150 160 170 180 190 200

ACGAGGCGCCACACGATGACCTCCCGGCGGCTGAGAGTCCCGGTGAGGTGAGCAGCAT

150 160 170 180 190 200 X

11. US-09-697-123B-22 (1-208)

US-09-697-123B-8 Sequence 8, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.50
Residue Identity = 87% Matches = 181 Mismatches = 27
Gaps = 0 Conservative Substitutions = 0

X
TCGAAGGAGAAAGGCTACGACCTGGCCCGGCGTGGTGGTATTAAGTCAACAAAGAGCTCGGCTGCATGTGC
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
10 20 30 40 50 60 70
100 110 120 130 140

CGAGCGCCATCAGCTGCTGACGCTGACCGGAGAGAGAGCTGCGCCACCATCGAATATCTGCTCCCTTGC
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
80 90 100 110 120 130 140
150 160 170 180 190 200

ACGAGGCTCAGACACGAGATGATCTCCGGCGGCGTGGAGTGGGAGAAACCGACGACAT
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
150 160 170 180 190 200

12. US-09-697-123B-22 (1-208)

US-09-697-123B-1 Sequence 1, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.50
Residue Identity = 87% Matches = 181 Mismatches = 27
Gaps = 0 Conservative Substitutions = 0

X
TCGAAGGAGAAAGGCTACGACCTGGCCCGGCGTGGTGGTATTAAGTCAACAAAGAGCTCGGCTGCATGTGC
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
10 20 30 40 50 60 70
100 110 120 130 140

CGAGCGCCATCAGCTGCTGACGCTGACCGGAGAGAGAGCTGCGCCACCATCGAATATCTGCTCCCTTGC
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
80 90 100 110 120 130 140
150 160 170 180 190 200

ACGAGGCTCAGACACGATGATCTCCGGCGGCGTGGAGTGGGAGAAACCGACGACAT
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
150 160 170 180 190 200

13. US-09-697-123B-22 (1-208)

US-09-697-123B-2 Sequence 2, Application US/09697123B

Initial Score = 180 Optimized Score = 180 Significance = 0.48
Residue Identity = 86% Matches = 180 Mismatches = 28
Gaps = 0 Conservative Substitutions = 0

X
TCGAAGGAGAAAGGCTACGACCTGGCCCGGCGTGGTGGTATTAAGTCAACAAAGAGCTCGGCTGCATGTGC
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
10 20 30 40 50 60 70
100 110 120 130 140

CGAGCGCCATCAGCTGCTGACGCTGACCGGAGAGAGAGCTGCGCCACCATCGAATATCTGCTCCCTTGC
TCAAGGAGAAAGGCTACGACCTGGCCCGGCGGCTGACAGGCTCAACAAAGAGAGCTCGGCTGCACGCTGC
X
80 90 100 110 120 130 140
150 160 170 180 190 200

80 90 100 110 120 130 140

ACGAGGCTCAGACCGCATGATGCTGTCCGGCGCGCTCGAGGTCCCGGTGGAACCGACGACAT
|||||
ACGAGGCTCAGACCGCATGATGCTGTCCGGCGCGCGCTCGAGGTCCCGGTGGAACCGACGACAT
150 160 170 180 190 200 X

14. US-09-697-123b-1 Sequence 17, Application US/09697123b

Initial Score = 177 Optimized Score = 177 Significance = 0.43
Residue Identity = 85% Matches = 177 Mismatches = 31
Gaps = 0 Conservative Substitutions = 0

10 20 30 40 50 60 70

TCAAGAGAACCGCTACGACCTGGCCCGCGCTGCGCTATTAAGTCAACAGAACGCTCGGCTGCATGTCG
|||||
TCAAGAGAACCGCTACGACCTGGCCCGCGCGCTGCGCTATTAAGTCAACAGAACGCTCGGCTGCATGTCG
150 160 170 180 190 200 X

80 90 100 110 120 130 140

GCGAGCCCATACGCTGCTGACCGCTGACCGAAGAGCTGCGCCACCATCGAATATCTGCTCGCTTGC
|||||
GCTCCCGATTCAGACGACCACTCTGACCGAAGAGAGCTGCGCCACCATCGAATATCTGCTCGCTTGC
150 160 170 180 190 200 X

150 160 170 180 190 200 X

ACGAGGCTCAGACCGCATGATGCTTCCGGCGCGCTCGAGGTCCCGGTGGAACCGACGACAT
|||||
ACGAGGCTCAGACCGCATGATGCTTCCGGCGCGCTCGAGGTCCCGGTGGAACCGACGACAT
150 160 170 180 190 200 X

15. US-09-697-123b-22 (1-208)

US-09-697-123b-2 Sequence 24, Application US/09697123b

Initial Score = 172 Optimized Score = 172 Significance = 0.33
Residue Identity = 82% Matches = 172 Mismatches = 36
Gaps = 0 Conservative Substitutions = 0

10 20 30 40 50 60 70

TCAAGAGAACCGCTACGACCTGGCCCGCGCTCGCTATTAAGTCAACAGAACGCTCGGCTGCATGTCG
|||||
TCAAGAGAACCGCTACGACCTGGCCCGCGCTCGCTATTAAGTCAACAGAACGCTCGGCTGCATGTCG
150 160 170 180 190 200 X

80 90 100 110 120 130 140

GCGAGCCCATACGCTGCTGACCGCTGACCGAAGAGCTGCGCCACCATCGAATATCTGCTCGCTTGC
|||||
AGAATGCGCCAACCCACCGACCGCTGACCGAAGAGAGCTGCGCCACCATCGAATATCTGCTCGCTTGC
150 160 170 180 190 200 X

150 160 170 180 190 200 X

ACGAGGCTCAGACCGCATGATGCTTCCGGCGCGCTCGAGGTCCCGGTGGAACCGACGACAT
|||||
ACGAGGCTCAGACCGCATGATGCTTCCGGCGCGCTCGAGGTCCCGGTGGAACCGACGACAT
150 160 170 180 190 200 X

16. US-09-697-123b-22 (1-208)

US-09-697-123b-2 Sequence 20, Application US/09697123b

Initial Score = 160 Optimized Score = 170 Significance = 0.10
Residue Identity = 84% Matches = 176 Mismatches = 29
Gaps = 3 Conservative Substitutions = 0

10 20 30 40 50 60 70

TCAAGAGAACCGCTACGACCTGGCCCGCGCTGCGCTATTAAGTCAACAGAACGCTCGGCTGCATGTCG
|||||
TCAAGAGAACCGCTACGACCTGGCCCGCGCTGCGCTATTAAGTCAACAGAACGCTCGGCTGCATGTCG
150 160 170 180 190 200 X

80 90 100 110 120 130 140

GCGAGCCCATACGCTGCTGACCGCTGACCGAAGAGCTGCGCCACCATCGAATATCTGCTCGCTTGC
|||||
GCGAGCCCATACGCTGCTGACCGCTGACCGAAGAGCTGCGCCACCATCGAATATCTGCTCGCTTGC
150 160 170 180 190 200 X

17. US-09-697-123b-1 Sequence 12, Application US/09697123b

Initial Score = 156 Optimized Score = 175 Significance = 0.02
Residue Identity = 86% Matches = 180 Mismatches = 27
Gaps = 1 Conservative Substitutions = 0

10 20 30 40 50 60 70

TCAAGAGAACCGCTACGACCTGGCCCGCGCTGCGCTATTAAGTCAACAGAACGCTCGGCTGCATGTCG
|||||
TCAAGAGAACCGCTACGACCTGGCCCGCGCTGCGCTATTAAGTCAACAGAACGCTCGGCTGCATGTCG
150 160 170 180 190 200 X

80 90 100 110 120 130 140

GCGAGCCCATACGCTGCTGACCGCTGACCGAAGAGCTGCGCCACCATCGAATATCTGCTCGCTTGC
|||||
GCGAGCCCATACGCTGCTGACCGCTGACCGAAGAGCTGCGCCACCATCGAATATCTGCTCGCTTGC
150 160 170 180 190 200 X

150 160 170 180 190 200 X

ACGAGGCTCAGACCGCATGATGCTTCCGGCGCGCTCGAGGTCCCGGTGGAACCGACGACAT
|||||
ACGAGGCTCAGACCGCATGATGCTTCCGGCGCGCTCGAGGTCCCGGTGGAACCGACGACAT
150 160 170 180 190 200 X

18. US-09-697-123b-22 (1-208)

US-09-697-123b-4 Sequence 4, Application US/09697123b

Initial Score = 153 Optimized Score = 178 Significance = -0.04
Residue Identity = 87% Matches = 183 Mismatches = 24
Gaps = 1 Conservative Substitutions = 0

10 20 30 40 50 60 70

TCAAGAGAACCGCTACGACCTGGCCCGCGCTGCGCTATTAAGTCAACAGAACGCTCGGCTGCATGTCG
|||||
TCAAGAGAACCGCTACGACCTGGCCCGCGCTGCGCTATTAAGTCAACAGAACGCTCGGCTGCATGTCG
150 160 170 180 190 200 X

80 90 100 110 120 130 140

GCGAGCCCATACGCTGCTGACCGCTGACCGAAGAGCTGCGCCACCATCGAATATCTGCTCGCTTGC
|||||
GCGATCCGATACCGCTGCTGACCGCTGACCGAAGAGCTGCGCCACCATCGAATATCTGCTCGCTTGC
150 160 170 180 190 200 X

150 160 170 180 190 200 X

ACGAGGCTCAGACCGCATGATGCTTCCGGCGCGCTCGAGGTCCCGGTGGAACCGACGACAT
|||||
ACGAGGCTCAGACCGCATGATGCTTCCGGCGCGCTCGAGGTCCCGGTGGAACCGACGACAT
150 160 170 180 190 200 X

19. US-09-697-123b-22 (1-208)

US-09-697-123b-9 Sequence 9, Application US/09697123b

Initial Score = 143 Optimized Score = 159 Significance = -0.23
Residue Identity = 80% Matches = 179 Mismatches = 29
Gaps = 15 Conservative Substitutions = 0

10 20 30 40 50 60 70

TCAAGAGAACCGCTACGACCTGGCCCGCGCTGCGCTATTAAGTCAACAGAACGCTCGGCTGCATGTCG
|||||
TCAAGAGAACCGCTACGACCTGGCCCGCGCTGCGCTATTAAGTCAACAGAACGCTCGGCTGCATGTCG
150 160 170 180 190 200 X

CGGAGCC-----CATCAGCTCGTCGACCGCTGACCGAAGAAGACGTCGTGGCCACCATCGAATATCTGTCC

[illegible]

> 0 <
0110 Intelligenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-697-123b-21.res made by shanley on Wed 13 Nov 102 14:11:41 PST.

Query sequence being compared: US-09-697-123B-21 (1-214)
Number of sequences searched: 26
Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123B-21 (1-214) with:
File: US09697123B.seq

```

100-
N
U 50-
M
B
E
R
O
F 10-
S
E 5-
O
U
E
N
C
S
S
0
SCORE 0 24 48 71 95 119 143 166 190 214
STDEV -2 -1 0 1 2

```

PARAMETERS

Similarity matrix Unitary 1 k-tuple 4
Mismatch penalty 5.00 Joining penalty 30
Gap penalty 0.33 Window size 205
Cutoff score 12
Randomization group 0

SEARCH STATISTICS

Scores: Mean 112 Median 112 Standard Deviation 42.66
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
Number of residues: 5077
Number of sequences searched: 26
Number of scores above cutoff: 25

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Init. Opt. Score	Sig. Frame
US-09-697-123B-21	US-09-697-123B-21 (1-214)	214	197	1.99

1. US-09-697-123B-2 Sequence 21, Application 214 214 214 2.39 0
The list of other best scores is:

Sequence Name	Description	Length	Init. Opt. Score	Sig. Frame
2. US-09-697-123B-7 Sequence 7, Application 214	1 standard deviation above mean ****	197	197	1.99
3. US-09-697-123B-1 Sequence 14, Application 214	1 standard deviation above mean ****	171	171	1.38
4. US-09-697-123B-1 Sequence 10, Application 208	0 standard deviation from mean ****	170	170	0.63
5. US-09-697-123B-1 Sequence 11, Application 208		139	139	0
6. US-09-697-123B-1 Sequence 19, Application 208		159	159	0.61
7. US-09-697-123B-8 Sequence 8, Application 208		125	125	0
8. US-09-697-123B-2 Sequence 23, Application 208		122	122	0.23
9. US-09-697-123B-1 Sequence 15, Application 208		117	117	0
10. US-09-697-123B-3 Sequence 3, Application 208		112	112	0.12
11. US-09-697-123B-2 Sequence 2, Application 208		112	112	0.00
12. US-09-697-123B-2 Sequence 24, Application 208		112	112	0.00
13. US-09-697-123B-1 Sequence 17, Application 208		111	111	0.00
14. US-09-697-123B-1 Sequence 18, Application 211		117	117	0.02
15. US-09-697-123B-2 Sequence 20, Application 205		110	110	0
16. US-09-697-123B-1 Sequence 16, Application 208		109	109	0.07
17. US-09-697-123B-6 Sequence 6, Application 208		109	109	0.07
18. US-09-697-123B-5 Sequence 5, Application 208		109	109	0.07
19. US-09-697-123B-1 Sequence 1, Application 208		109	109	0.07
20. US-09-697-123B-2 Sequence 12, Application 208		109	109	0.07
21. US-09-697-123B-1 Sequence 22, Application 207		108	108	0.14
22. US-09-697-123B-9 Sequence 9, Application 207		106	106	0
23. US-09-697-123B-4 Sequence 4, Application 223		82	82	0.68
24. US-09-697-123B-1 Sequence 13, Application 223		81	81	0.73
25. US-09-697-123B-2 Sequence 25, Application 19	2 standard deviations below mean ****	19	19	-2.18

1. US-09-697-123B-21 (1-214)

US-09-697-123B-2 Sequence 21, Application US/09697123B

Initial Score - 214 Optimized Score - 214 Significance - 2.39
Residue Identity - 100% Matches - 214 Mismatches - 0
Gaps - 0 Conservative Substitutions - 0

X
TCAAGGAGAGCGCTACGATCTGCGCCCGCGTGGTGTGACAGGTGAACAAGAGTGGCGCGCGCA
|||||
TCAAGGAGAGCGCTACGATCTGCGCCCGCGTGGTGTGACAGGTGAACAAGAGTGGCGCGCGCA
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
CCAATCCGCTCAGTGTGACACACCCCTACCGAGAGAGAGTGTGCGCACACATCGATCTGTGTC
|||||
CCAATCCGCTCAGTGTGACACACCCCTACCGAGAGAGAGTGTGCGCACACATCGATCTGTGTC
80 90 100 110 120 130 140
GCTGTGACAGGCGCGTGTGACACACCCCTACCGAGAGAGAGTGTGCGCACACATCGATCTGTGTC
150 160 170 180 190 200 210 X
GCTGTGACAGGCGCGTGTGACACACCCCTACCGAGAGAGAGTGTGCGCACACATCGATCTGTGTC
150 160 170 180 190 200 210 X

2. US-09-697-123B-21 (1-214)

US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score - 197 Optimized Score - 197 Significance - 1.99
Residue Identity - 92% Matches - 197 Mismatches - 17
Gaps - 0 Conservative Substitutions - 0

X
TCAAGGAGAGCGCTACGATCTGCGCCCGCGTGGTGTGACAGGTGAACAAGAGTGGCGCGCGCA
|||||
TCAAGGAGAGCGCTACGATCTGCGCCCGCGTGGTGTGACAGGTGAACAAGAGTGGCGCGCGCA
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
CCAATCCGCTCAGTGTGACACACCCCTACCGAGAGAGAGTGTGCGCACACATCGATCTGTGTC
|||||
CCAATCCGCTCAGTGTGACACACCCCTACCGAGAGAGAGTGTGCGCACACATCGATCTGTGTC
80 90 100 110 120 130 140
GCTGTGACAGGCGCGTGTGACACACCCCTACCGAGAGAGAGTGTGCGCACACATCGATCTGTGTC
150 160 170 180 190 200 210 X
GCTGTGACAGGCGCGTGTGACACACCCCTACCGAGAGAGAGTGTGCGCACACATCGATCTGTGTC
150 160 170 180 190 200 210 X

8. US-09-697-1238-21 (1-214)

CGAATCCGGCTAGGTGACCAACCAACCCCTCAACCGAGGAGACGTCGTCGCACCAATCGATACCGGCG
 80 90 100 110 120 130 140
 CCAATCCGGCTAGGTGACCAACCAACCCCTCAACCGAGGAGACGTCGTCGCACCAATCGATACCGGCG
 150 160 170 180 190 200 210
 ACACCCGTCGCCCATCATCAGACGACCACTTGACCCAGAGGAGACGTCGTCGCGACCAATCGATACCGGCG

	70	80	90	100	110	120	130
150		160	170	180	190	200	210
	CCTGTACAGCAGGCGGAGC	CAACGATGTACCGCGCGCGCGCGCTG	CCAGGTGCGGTGATGTGGCAGCAT				X
140	GCCTGCACGAGGCGCAC	CAACGATGTACCGCGCGCGAGT	CCAGGTGCGGTGGAACCGCAGCAT				
	150	160	170	180	190	200	X

14. US-09-697-123B-21 (1-214)
US-09-697-123B-1 Sequence 18, Application US/09697123B

Initial Score	=	111	Optimized Score	=	170	Significance	=	-0.02
Residue Identity	=	82%	Matches	=	176	Mismatches	=	35
Gaps	=	3	Conservative Substitutions	=			=	0

[illegible]

GCCTTCACGAGGGCCAGCAGATGACCGCCCCCGCGCGTGCAGATCCCGTGATGTGACAGACA
150 160 170 180 190 200 210

GCCCTCATCAGGCCAACAAGCATGACCCTCGCGGTGAGTCGAGTCCCGCTGAGAGTTCGACGACA
150 160 170 180 190 200 210

15. US-09-697-123B-21 (1-214)
US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score	-	109	Optimized Score	-	159	Significance	-	-0.07
Residue Identity	-	79%	Matches	-	171	Mismatches	-	34
Gaps	-	9	Conservative Substitutions	-			-	0

X	10	20	30	40	50	60	70
TCACGAGAACCGCTACGATCTGGCCCGCGGGTCCGTTCAAGTGAACAGAAAGACTGGGCTGGGGCGCA							
TCACGAGAACCGCTACGACCTGGCGGGTGTCCGGCCCTTCAAGTCAACAGAAAGACTGGGCTGCAGCGGG							
X	10	20	30	40	50	60	70

CCAAATCGGCTCAGTGTACGCACCACCAACCCTCACCGAGAGACGGTGTGGCACCATGTACTACTGTGC
80 90 100 110 120 130 140
| | | | | | | | | | | | | | | |
GCGAGCC-----GATCACCGCTCGACCTGACCCGAGAGACGGTGTGGCACCATGTACTACTGTGC
80 90 100 110 120 130

GCCTGCACGAGGGCCACACCATGATACCGCCCGCGCGCGCTGAGGTGCGGTGATGTGACGACAT	150	160	170	180	190	200	210	X
GCCTGCACGAGGGCCACACCATGATACCGCCCGCGCGCGCTGAGGTGCGGTGATGTGACGACAT	150	160	170	180	190	200	210	X
GCCTGCACGAGGGCCACACCATGATACCGCCCGCGCGCGCTGAGGTGCGGTGATGTGACGACAT	150	160	170	180	190	200	210	X
GCCTGCACGAGGGCCACACCATGATACCGCCCGCGCGCGCTGAGGTGCGGTGATGTGACGACAT	150	160	170	180	190	200	210	X

16. US-09-697-123B-21 (1-214)
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score	=	109	Optimized Score	=	162	Significance	=	-0.07
Residue Identity	=	78%	Matches	=	169	Mismatches	=	39
Gaps	=	6	Conservative Substitutions	=	0			

X 10 20 30 40 50 60 70
TCAAGCAAGACCGCTACGATCTGGCCGCTGGTCCGTAACAAGTGAACAAGACCTGGGCTGGGCGGCA
|||||
TCAAGGAAAGACCGCTACGACCTGCGCGCTCGGTGCTAATAGGTCACACAGAAGACCTGGGGCTGCATGTGG
10 20 30 40 50 60 70

CCAATCCGGCTCAGGTGACCAACCACCCCTACCGGAGAAGTGTGTGCCAACCATTGAGTAACCTGTGTC
 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 GCGAGCC-----CATCACGTGTGTGAGCGCTGACCGAAGAAGACGTGTGTGCCACCATCGAAATATCTGTGTC

150 160 170 180 190 200 210
GCTCTGACGACGAGGCGACGACCAAGTGAACCCCGCCGCGCGCGTCCAGAGTCCGCGTGAATGTGCACGACAT
|||||
GCTTTCGACGAGGCTGAGACCAACGATGACGCTTCCGCGGCGCGCTGACGCTGCGCGGTGGAACCGACGACAT
|||||
140 150 160 170 180 190 200
X

17. US-09-697-123B-21 (1-214)
US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score	=	109	Optimized Score	=	162	Significance	=	-0.07
Residue Identity	=	78%	Matches	=	169	Mismatches	=	39
Gaps	=	6	Conservative Substitutions	=	0			

X 10 20 30 40 50 60 70
TCAGAGAGAGGCGCTACGATCTCGCCCGCGGGTCGGTACAAAGTGAAACAAAGAGCTGGCCGCGGCA
|||||
TCAGAGAGAGGCGCTACGACCTCGGCCCGCGTGGCTCTTAAGGTCAACAAAGAGCTGGCGCTCATGTGCG
X 10 20 30 40 50 60 70

CCAAATCGGCTAGTGTACCAACCAACCCTCACCGAAGAACCTCGTCGCCACCATGCAGTAACCTGGT
| | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | |
GGAGACC-----CATACGTCGTCGAACGTGACCGAAGAACACTCTGGCCACATCAATATCTGGT

[illegible]

18. US-09-697-123B-21 (1-214)
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial score	=	109	Optimized Score	=	162	Significance	=	-0.07
Residue Identity	=	78	Matches	=	169	Mismatches	=	39
Gaps	=	6	Conservative Substitutions	=			=	0

X
10 20 30 40 50 60 70
TAAAGAGAAAGCGTACATTTGGCCCGCTGGTGCAGTGAACAAGAAAGCTGGCGTGGCGGCA
|||||
TAAAGAGAGCGCTACACTTGGCCCGCTGGTGCCTTAAGGTCAACAAGAAAGCTGGCGTCAATGTG
X 10 20 30 40 50 60 70

CCAAATCGGCTCAGGTGCACCAACCCTCACCGGGAAGAAGTGTGGCCACCATTCGAGTAACCTGTGCT
80 90 100 110 120 130 140
| | | | |
GGGAGCC-----CATCAGCTGCTGCACGCTGACCCGAGAAGACGTTGGGCCACCATTCGAAATTATCTGTGTC
80 90 100 110 120 130

	150	160	170	180	190	200	210	X
GCCTGCAGAGAGGGCCAGACAGATATACCGGGCCCCGGCGGCGTGAAGTCCGGTGAATGTGTACAGACAT								
GCCTTCACAGAGGGCTACAGCCAGCATATACCTCTCGCGGGCGGGCTGAAGTCCCGGTGGAAACCGACACAT								
	150	160	170	180	190	200	X	

19. US-09-697-123B-21 (1-214)
US-09-697-123B-1 Sequence 1, Application US/09697123B

Initial Score =	109	Optimized Score =	170	Significance =	-0.07
Residue Identity =	82%	Matches =	177	Mismatches =	31
Gaps =	6	Conservative Substitutions			0

X 10 20 30 40 50 60 70
TCACGAGAACGCTACGATCTGCCCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCTGGCGCA
|||||
TCACGAGAACGCTACGATCTGCCCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCTGGCGCA
X 10 20 30 40 50 60 70
CGAATCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
|||||
CGAATCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
X 10 20 30 40 50 60 70
GCGATCC-----GATCACCGATTGCGACGCTGACCGAAGAGATGTCGTCCACCATCGATCGGTTC
80 90 100 110 120 130 140
GCGTCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
150 160 170 180 190 200 210 X
GCGTCCGCGAGGCGCAGACCAAGATGACCGCCCGCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCA
|||||
GCGTCCGAGAGGCGCAGACCAAGATGACCGCCCGCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCA
140 150 160 170 180 190 200 X

20. US-09-697-123B-21 (1-214)
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 108 Optimized Score = 161 Significance = -0.09
Residue Identity = 78% Matches = 168 Mismatches = 40
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAACGCTACGATCTGCCCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCTGGCGCA
|||||
TCAAGGAGAACGCTACGATCTGCCCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCTGGCGCA
X 10 20 30 40 50 60 70
CGAATCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
|||||
CGAATCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
X 10 20 30 40 50 60 70
GCGATCC-----GATCACCGATTGCGACGCTGACCGAAGAGATGTCGTCCACCATCGATCGGTTC
80 90 100 110 120 130 140
GCGTCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
150 160 170 180 190 200 210 X
GCGTCCGAGAGGCGCAGACCAAGATGACCGCCCGCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCA
|||||
GCGTCCGAGAGGCGCAGACCAAGATGACCGCCCGCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCA
140 150 160 170 180 190 200 X

21. US-09-697-123B-21 (1-214)
US-09-697-123B-1 Sequence 12, Application US/09697123B

Initial Score = 106 Optimized Score = 158 Significance = -0.14
Residue Identity = 79% Matches = 170 Mismatches = 37
Gaps = 7 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAACGCTACGATCTGCCCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCTGGCGCA
|||||
TCAAGGAGAACGCTACGATCTGCCCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCTGGCGCA
X 10 20 30 40 50 60 70
CGAATCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
|||||
CGAATCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
X 10 20 30 40 50 60 70
GCGATCC-----GATCACCGATTGCGACGCTGACCGAAGAGATGTCGTCCACCATCGATCGGTTC
80 90 100 110 120 130 140
GCGTCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
150 160 170 180 190 200 210 X
GCGTCCGAGAGGCGCAGACCAAGATGACCGCCCGCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCA
|||||
GCGTCCGAGAGGCGCAGACCAAGATGACCGCCCGCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCA
140 150 160 170 180 190 200 X

22. US-09-697-123B-21 (1-214)
US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 83 Optimized Score = 152 Significance = -0.68
Residue Identity = 73% Matches = 168 Mismatches = 40
Gaps = 21 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAACGCTACGATCTGCCCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCTGGCGCA
|||||
TCAAGGAGAACGCTACGATCTGCCCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCTGGCGCA
X 10 20 30 40 50 60 70
CGAATCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
|||||
CGAATCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
X 10 20 30 40 50 60 70
GCGATCC-----GATCACCGATTGCGACGCTGACCGAAGAGATGTCGTCCACCATCGATCGGTTC
80 90 100 110 120 130 140
GCGTCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
150 160 170 180 190 200 210

23. US-09-697-123B-21 (1-214)
US-09-697-123B-4 Sequence 4, Application US/09697123B

Initial Score = 82 Optimized Score = 157 Significance = -0.70
Residue Identity = 78% Matches = 169 Mismatches = 38
Gaps = 7 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAACGCTACGATCTGCCCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCTGGCGCA
|||||
TCAAGGAGAACGCTACGATCTGCCCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCTGGCGCA
X 10 20 30 40 50 60 70
CGAATCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
|||||
CGAATCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
X 10 20 30 40 50 60 70
GCGATCC-----GATCACCGATTGCGACGCTGACCGAAGAGATGTCGTCCACCATCGATCGGTTC
80 90 100 110 120 130 140
GCGTCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
150 160 170 180 190 200 210

24. US-09-697-123B-21 (1-214)
US-09-697-123B-1 Sequence 13, Application US/09697123B

Initial Score = 81 Optimized Score = 155 Significance = -0.73
Residue Identity = 76% Matches = 176 Mismatches = 32
Gaps = 21 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAACGCTACGATCTGCCCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCTGGCGCA
|||||
TCAAGGAGAACGCTACGATCTGCCCGGCTGCGTCAAGGTGAAGCAAGCTGGCGCTGGCGCA
X 10 20 30 40 50 60 70
CGAATCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
|||||
CGAATCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
X 10 20 30 40 50 60 70
GCGATCC-----GATCACCGATTGCGACGCTGACCGAAGAGATGTCGTCCACCATCGATCGGTTC
80 90 100 110 120 130 140
GCGTCCGCGCTCAGGTGACACACACCTCAGCGAGAAAGCTGCGGCACTGAGTGTACTGTGTC
150 160 170 180 190 200 210

```

      80      90      100      110      120      130
150  GCCTGCAC-----GAGGGCAGACCAC---GATGACCCCGCCGCGGCTCGAGGTCCGGTGG
      |||||||
140  GCCTGCACACGCGCTCTCAAGGTGGCCAGGCCCGCCGCTTATGACTGTCCCGCGGGTTCGAGGTCCGGTGG
      150      160      170      180      190      200      210
210  ATGTGGACGACAT
      | |||||||
      AACCGACGACAT
220  X

```

25. US-09-697-123b-21 (1-214)
US-09-697-123b-2 Sequence 25, Application US/09697123B

```

Initial Score      = 19  Optimized Score = 19  Significance = -2.18
Residue Identity = 100%  Matches      = 19  Mismatches    = 0
Gaps              = 0    Conservative Substitutions = 0
X
TCAAGCAGAAAGCGCTACGA
|||||
TCAAGCAGAAAGCGCTACGA
X
10

```

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig. Frame
---------------	-------------	--------	-------------	------------	------------

X 10 20 30 40 50 60 70
 TC AAGG AAGC GCGT AC GAC CTTGG CCGGTG CGCCGT AC AAGTCA CAAGAAG CTTGG CCGTCA CAGCGG
 |||||
 TC AAGG AAGC GCGT AC GAC CTTGG CCGGTG CGCCGT AC AAGTCA CAAGAAG CTTGG CCGTCA CAGCGG
 |||||
 TC AAGG AAGC GCGT AC GAC CTTGG CCGGTG CGCCGT AC AAGTCA CAAGAAG CTTGG CCGTCA CAGCGG
 |||||
 X 10 20 30 40 50 60 70
 TC AAGG AAGC GCGT AC GAC CTTGG CCGGTG CGCCGT AC AAGTCA CAAGAAG CTTGG CCGTCA CAGCGG
 |||||
 TC AAGG AAGC GCGT AC GAC CTTGG CCGGTG CGCCGT AC AAGTCA CAAGAAG CTTGG CCGTCA CAGCGG
 |||||
 TC AAGG AAGC GCGT AC GAC CTTGG CCGGTG CGCCGT AC AAGTCA CAAGAAG CTTGG CCGTCA CAGCGG
 |||||

8. US-09-697-123B-20 (1-205)
US-09-697-123B-8 Sequence 8, Application US/09697123B

Initial Score	=	163	Optimized Score	=	173	Significance	=	0.49
Residue Identity	=	86%	Matches	=	179	Mismatches	=	26
Gaps	=	3	Conservative Substitutions	=			=	0

X 10 20 30 40 50 60 70
 TCACGAGAACCGTAGCAGCTTGGCGGTGTGGCGGCTCAAGGTCACACAGAACCTTGGGCTCAGACGGG
 |||||
 TCACGAGAACCGTAGCAGCTTGGCGGTGTGGCGGCTCAAGGTCACACAGAACCTTGGGCTCAGACGGG
 X 10 20 30 40 50 60 70

CGGAGCCGATATACCAAGCTCGACGCTGACCGAGAAGACGTCGTGGCCACCATCGATGATAGCTGCGCTGCG
80 90 100 110 120 130 140
ATCATCCGATCACCACACGACGCTGACCGAGAAGACGTCGTGGCCACCATCGATGATAGCTGCGCTGCG
90 100 110 120 130 140

ACGAGGGGACGACCATATACGCTGCC---CGGCATCGAGTGGCCGCTGGAGACCGACGACAT
150 160 170 180 190 200
ACGAGGGCCAGCGCACCATACCGTCCGCGGCGGAGGTCTAGGTGGAGAAACCGACGACAT
150 160 170 180 190 200

9. US-09-697-123B-20 (1-205)
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score	=	161	Optimized Score	=	171	Significance	=	0.45
Residue Identity	=	85%	Matches	=	177	Mismatches	=	28
Gaps	=	3	Conservative Substitutions	=			=	0

X 10 20 30 40 50 60 70
TCAAGAGCAACGGCTACACACTGCGCGCTGTGGCCCTACAAAGTCAACAGAGCTCGGCTGCACCC
|||||
TCAAGAGGAGCCCTACGACCTGCGCGCGCTGTGGCGCTAATAAGTCAACAGAGAGCTCGGCTGCATGT
X 10 20 30 40 50 60 70

GCAGCCGATCACCACTGTGACGTGACCGAAGAAGTGTGCCACCATTGGTAGCCT
80 90 100 110 120 130 140
TTTTTTT TTTTTTT TTTTTTT TTTTTTT TTTTTTT TTTTTTT
GCGAGCCATCACGTGTGACGTGACCGAAGAAGTGTGCCACCATTATGTGCGCTT
80 90 100 110 120 130 140

[illegible]

10. US-09-697-123B-20 (1-205)
US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score	=	161	Optimized Score	=	171	Significance	=	0.45
Residue Identity	=	85%	Matches	=	177	Matches	=	28
Gaps	=	3	Conservative Substitutions	=	0	Mismatches	=	0

X 10 20 30 40 50 60 70
 TCACGAGGAGGACCGTACGACGTCGGCGTGTGCGCCGCTACAAAGTCAACAAGCTGGCGCTGACACGG
 TCACGAGGAGGACCGTACGACGTCGGCGTGTGCGCCGCTACAAAGTCAACAAGCTGGCGCTGACACGG
 X 10 20 30 40 50 60 70
 TCACGAGGAGGACCGTACGACGTCGGCGTGTGCGCCGCTACAAAGTCAACAAGCTGGCGCTGACACGG

[illegible][illegible]

11. US-09-697-123B-20 (1-205)
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score	-	161	Optimized Score	-	171	Significance	-	0.45
Residue Identity	-	85%	Matches	-	177	Mismatches	-	28
Gaps	-	3	Conservative Substitutions	-			-	0

X 10 20 30 40 50 60 70
 TCATAGGAGAACGGCTACGAGACTGTGGCGCTGTGCGCCGCTCAAGTCAACAAGAGAGCTCGGCTGCACCGG
 TCATAGGAGAACGGCTACGAGACTGTGGCGCTGTGCGCCGCTCAAGTCAACAAGAGAGCTCGGCTGCACCGG
 TCATAGGAGAACGGCTACGAGACTGTGGCGCTGTGCGCCGCTCAAGTCAACAAGAGAGCTCGGCTGCACCGG
 X 10 20 30 40 50 60 70

GCGAGCGATCCACCACTCTGAGCTGACCGGAGAAACGTCGCGACACATCGAGATCTTGGTGGCCTGG
 80 90 100 110 120 130 140
 GCGAGCCATCAGCTGTCGAGACCTGACCGAAGACGCTGGCCACACATCAATATGTGGTGGCTGGC
 90 100 110 120 130 140

[illegible]

12. US-09-697-123B-20 (1-205)
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score	-	160	Optimized Score	-	170	Significance	-	0.422
Residue Identity	-	84%	Matches	-	176	Mismatches	-	29
Gaps	-	3	Conservative Substitutions	-			-	0

A
10 20 30 40 50 60 70
TCAAGSAGAACGCGTACGAACCTGGCGGTGTCGCCCGCTACAAAGTCAACAAGAGCTGCCTGCACCCGGC
|||||
TCAAGSAGAAGCGCTACGAACCTGGCGGTGCTATTAAGGTCAACAAGAGCTGCCTGCATCTCC
X 10 20 30 40 50 60 70

GCGAGCGATCACACGCTCGACGCTGACCGGAAAGCGCCGCGACATCGATCGGTGGCGCTGG
 80 90 100 110 120 130 140
 GCGAGCCCATCAGTCGTGAAAGCGGAAACGATCGTGGCCACATCGAATATATGGTCGGCTTGG
 80 90 100 110 120 130 140

ACAGGGCCACGCCCACGATGACCGTCCC---CGCATGAGGTGGCGGTGGAGACCCAGCAGCA
150 160 170 180 190 200
ACAGGGTTCAGACCCACGATGATGTTCCCGGCGCGGTGAGGTGGCGGTGGAACCCAGCAGCA
150 160 170 180 190 200

13. US-09-697-123B-20 (1-205)
US-09-697-123B-1 Sequence 17, Application US/09697123B

Initial Score	=	160	Optimized Score	=	170	Significance	=	0.42
Residue Identity	=	84%	Matches	=	176	Mismatches	=	29
Gaps	=	3	Conservative Substitutions	=			=	0

TCAAGAGAGAAACGGGTACGACCTGTGCGCGCTACAGGTCAACAGACGTGGCTGCACCC
 10 20 30 40 50 60 70
 TCAAGAGAGAAACGGGTACGACCTGTGCGCGCTACAGGTCAACAGACGTGGCTGCACCC
 10 20 30 40 50 60 70
 TCAAGAGAGAAACGGGTACGACCTGTGCGCGCTACAGGTCAACAGACGTGGCTGCACCC
 10 20 30 40 50 60 70

[illegible]

X
 10
 20
 30
 40
 50
 60
 70
 TCAGGAGAACGGCTACGACCTGGCGCTGTCGGCGCTACAGGATCAACAGAAGCTGGCGCTGCACGGG

```
|||||
TCAGAGAGAGGCGCTACACCTGGCCCGCTGGCCGCTACAAAGTCAACAAGAGCTGGCTGAACACCG
X 10 20 30 40 50 60 70
80 90 100 110 120 130 140
GGGACCCGATACACAGCTGACCGTGAACGAGAGAGCTGTCGCCACCATGAGTACCTGGTGGCTTC
ATCATCGCATACACACCGACGCGTGAACGAGAGAGCTGTCGCCACCATGAGTACCTGGTGGCTTC
80 90 100 110 120 130 140
150 160 170 180 190 200 210
ACGAGGCC-----CAGCCACGATGACCTGCCCGC---ATCGAGTGGCCGGTGGAGACCG
|||
ACCATCGCTCTCAGGGTGGCCAGGCCCGCTTATGACTGTCCCGGGGGTGGAGTCCCGGGTGAACCG
150 160 170 180 190 200 210
200 X
ACGACAT
|||||
ACGACAT
220 X
```

20. US-09-697-123b-20 (1-205)
US-09-697-123b-9 Sequence 9, Application US/09697123b

```
Initial Score = 131 Optimized Score = 163 Significance = -0.22
Residue Identity = 82% Matches = 183 Mismatches = 22
Gaps = 18 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGGCGCTACACCTGGCCCGCTGTCGCGCTTCAAGTCAACAAGAGCTGGCTGGACCGG
|||||
TCAAGAGAGAGGCGCTACACCTGGCCCGCTGTCGCGCTTCAAGTCAACAAGAGCTGGCTGGACCGG
X 10 20 30 40 50 60 70
80 90 100 110 120 130 140
GGGACCCGATACACAGCTGACCGTGAACGAGAGAGCTGTCGCCACCATGAGTACCTGGTGGCTTC
|||||
GGGACCCGATACCTGCTCCACCTGACCGAGAGAGAGCTGTCGCCACCATGAGTACCTGGTGGCTTC
80 90 100 110 120 130 140
150 160 170 180 190 200 210
AC-----GAGGGCGAGCCAC---GATGACCTGCC---CGGATGCGAGTGGCCGCTGGAGACCG
|||
ACACGCCCGCTAGCGATGCGCAGCCCGCTCATGACTGTCCCGGGGATCGAGTGGCCGCTGGAGACCG
150 160 170 180 190 200 210
200 X
ACGACAT
|||||
ACGACAT
220 X
```

21. US-09-697-123b-20 (1-205)
US-09-697-123b-1 Sequence 14, Application US/09697123b

```
Initial Score = 122 Optimized Score = 162 Significance = -0.42
Residue Identity = 81% Matches = 174 Mismatches = 31
Gaps = 9 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGGCGCTACACCTGGCCCGCTGTCGCGCTTCAAGTCAACAAGAGCTGGCTGGACCGG
|||||
TCAAGAGAGAGGCGCTACACCTGGCCCGCTGTCGCGCTTCAAGTCAACAAGAGCTGGCTGGACCGG
X 10 20 30 40 50 60 70
80 90 100 110 120 130 140
GGGACCCG-----ATCAGCAGCTGACCGCTGACCGAGAGAGAGCTGTCGCCACCATGAGTACCTGGCTTC
|||||
CCGAGTGGCGCTGACCGCTGACCGCTGACCGAGAGAGCTGTCGCCACCATGAGTACCTGGCTTC
80 90 100 110 120 130 140
```

```
140 150 160 170 180 190 200 X
GCTTCACAGAGGCGCCACCGCATGACCGTCCC---CGGCATGAGTGGCCCGTGGAGACCGACACAT
|||||
GCTTCACAGAGGCGCCACCGCATGAGTGGTTCGCCGCGCGCTGAGAGTGGCTGGTGGAGACCGACAT
150 160 170 180 190 200 210 X
22. US-09-697-123b-20 (1-205)  
US-09-697-123b-2 Sequence 21, Application US/09697123b
```

```
Initial Score = 109 Optimized Score = 159 Significance = -0.71
Residue Identity = 79% Matches = 171 Mismatches = 34
Gaps = 9 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGGCGCTACACCTGGCCCGCTGTCGCGCTTCAAGTCAACAAGAGCTGGCTGGACCGG
|||||
TCAAGAGAGAGGCGCTACACCTGGCCCGCTGTCGCGCTTCAAGTCAACAAGAGCTGGCTGGACCGG
X 10 20 30 40 50 60 70
```

```
GGGAGCC-----GATCAGCAGCTGACCGCTGACCGAGAGAGAGCTGTCGCCACCATGAGTACCTGGCTTC
|||||
CCAAATCCGGCTCAGGTGACCGACCGACCGCTTCAAGTCAACAAGAGCTGGCTGGACCGGCTTC
80 90 100 110 120 130 140
```

```
140 150 160 170 180 190 200 X
GCTTCACAGAGGCGCCACCGCATGACCGTCCC---CGGCATGAGTGGCCCGTGGAGACCGACAT
|||||
GCTTCACAGAGGCGCCACCGCATGAGTGGTTCGCCGCGCGCTGAGAGTGGCTGGAGTGGAGTGGAGT
150 160 170 180 190 200 210 X
```

23. US-09-697-123b-20 (1-205)
US-09-697-123b-7 Sequence 7, Application US/09697123b

```
Initial Score = 109 Optimized Score = 158 Significance = -0.71
Residue Identity = 79% Matches = 170 Mismatches = 35
Gaps = 9 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGGCGCTACACCTGGCCCGCTGTCGCGCTTCAAGTCAACAAGAGCTGGCTGGACCGG
|||||
TCAAGAGAGAGGCGCTACACCTGGCCCGCTGTCGCGCTTCAAGTCAACAAGAGCTGGCTGGACCGG
X 10 20 30 40 50 60 70
```

```
GGGAGCC-----GATCAGCAGCTGACCGCTGACCGAGAGAGAGCTGTCGCCACCATGAGTACCTGGCTTC
|||||
CCAAATCCGGCTGAGTACTGCGACCGACCGCTTCAAGTCAACAAGAGCTGGCTGGACCGGCTTC
80 90 100 110 120 130 140
```

```
140 150 160 170 180 190 200 X
GCTTCACAGAGGCGCCACCGCATGACCGTCCC---CGGCATGAGTGGCCCGTGGAGACCGACAT
|||||
GCTTCACAGAGGCGCCACCGCATGACCGTCCC---CGGCATGAGTGGCCCGTGGAGACCGACAT
150 160 170 180 190 200 210 X
```

24. US-09-697-123b-20 (1-205)
US-09-697-123b-1 Sequence 18, Application US/09697123b

```
Initial Score = 108 Optimized Score = 152 Significance = -0.74
Residue Identity = 77% Matches = 163 Mismatches = 42
Gaps = 6 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGGCGCTACACCTGGCCCGCTGTCGCGCTTCAAGTCAACAAGAGCTGGCTGGACCGG
|||||
TCAAGAGAGAGGCGCTACACCTGGCCCGCTGTCGCGCTTCAAGTCAACAAGAGCTGGCTGGACCGG
X 10 20 30 40 50 60 70
```


FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Query sequence being compared:	US-09-697-123B-19 (1-208)
Number of sequences searched:	26
Number of scores above cutoff:	25

Results of the initial comparison of US-09-697-123B-19 (1-208) with
File : US09697123B.seq

100-
-
N
U
M
B
E
R
O
F
S
E
O
U
E
N
C
E
S
0
-
*
23
-2
46
69
-1
92
116
0
139
162
185
1
208
SCORE
STDEV

Similarity matrix	Unary	4
Mismatch penalty	1	30
Gap penalty	5.00	207
Gap size penalty	0.33	
Cutoff score	12	
Randomization group	0	

Scores:	Mean	Median	Standard Deviation
	154	177	48.92
Times:	CPU		Total Elapsed
	00:00:00.00		00:00:01.00

Number of residues:	5077
Number of sequences searched:	26
Number of scores above cutoff:	25

The scores below are sorted by initial score. Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig. Frame
-----	-----	-----	-----	-----	-----

Sequence Name	Description	Length	Score	Score	Sig.	Frame
---------------	-------------	--------	-------	-------	------	-------

25.	US-09-697-123B-2	Sequence 25,	Application	19	19	19	-2.76	0
24.	US-09-697-123B-7	Sequence 7,	Application	214	115	177	-0.80	0
23.	US-09-697-123B-1	Sequence 1,	Application	214	116	168	-0.78	0
22.	US-09-697-123B-2	Sequence 11,	Application	211	125	170	-0.59	0
21.	US-09-697-123B-1	Sequence 18,	Application	223	143	169	-0.22	0
20.	US-09-697-123B-9	Sequence 9,	Application	223	146	173	-0.16	0
19.	US-09-697-123B-4	Sequence 4,	Application	207	147	173	-0.14	0
18.	US-09-697-123B-1	Sequence 12,	Application	207	157	176	-0.06	0
17.	US-09-697-123B-2	Sequence 20,	Application	208	175	175	0.43	0
16.	US-09-697-123B-2	Sequence 24,	Application	208	176	176	0.45	0
15.	US-09-697-123B-2	Sequence 22,	Application	208	180	180	0.53	0
14.	US-09-697-123B-2	Sequence 3,	Application	208	181	181	0.55	0
13.	US-09-697-123B-3	Sequence 3,	Application	208	182	182	0.57	0
12.	US-09-697-123B-2	Sequence 5,	Application	208	182	182	0.57	0
11.	US-09-697-123B-6	Sequence 6,	Application	208	182	182	0.57	0
10.	US-09-697-123B-1	Sequence 16,	Application	208	182	182	0.57	0
9.	US-09-697-123B-1	Sequence 8,	Application	208	183	183	0.59	0
8.	US-09-697-123B-8	Sequence 17,	Application	208	183	183	0.59	0
7.	US-09-697-123B-1	Sequence 1,	Application	208	184	184	0.61	0
6.	US-09-697-123B-1	Sequence 2,	Application	208	184	184	0.61	0
5.	US-09-697-123B-2	Sequence 10,	Application	208	186	186	0.65	0
4.	US-09-697-123B-1	Sequence 11,	Application	208	187	187	0.67	0
3.	US-09-697-123B-1	Sequence 15,	Application	208	190	190	0.74	0
2.	US-09-697-123B-1	Sequence 11,	Application	208	190	190	0.74	0
1.	US-09-697-123B-1	Sequence 11,	Application	208	190	190	0.74	0

Initial Score	-	208	Optimized Score	-	208	Significance	-	1.10
Residue Identity	-	100%	Matches	-	208	Mismatches	-	0
Gaps	-	0	Conservative Substitutions	-			-	0

TCGAAGAGAAAGCGCTACGACCTGGCCCGCCGCTACAAAGTCAACAAGAGCTGGGGCTGAAGCCCGC
X 10 20 30 40 50 60 70
TCGAAGAGAAAGCGCTACGACCTGGCCCGCCGCTACAAAGTCAACAAGAGCTGGGGCTGAAGCCCGC
X 10 20 30 40 50 60 70

GGCAGGCCATCGTCGTCGACTGTGACCCAGAAACACTGTCGCCACATCGAGTACTGCTGGCGCTGT
 80
 GCCAGGCCATCGTCGTCGACTGTGACCCAGAAACACTGTCGCCACATCGAGTACTGCTGGCGCTGT
 90
 100
 110
 120
 130
 140

150 160 170 180 190 200
 ACAGAGGCCAGACACGATGACCGTCCCGGGGGGTGAGTCCCGGTGAGGTGACGACATAT
 150 160 170 180 190 200
 ACAGAGGCCAGACACGATGACCGTCCCGGGGGGTGAGTCCCGGTGAGGTGACGACATAT

2. US-09-697-1238-19 (1-208)
US-09-697-1238-1 Sequence 15, Application US/096971238E

Conservative Substitutions	-	0
Matches	-	18
Mismatches	-	0
Significance	-	0.74
Optimized Score	-	190
91% Matches	-	190
Identity	-	190
Residue	-	190
Wildcards	-	0

[illegible]

8. US-09-697-123B-19 (1-208)
US-09-697-123B-8 Sequence 8, Application US/09697123B

X
10 20 30 40 50 60 70
TCAAGGAGAGCGCTACGACTGGCCCGCTGGCCCTACAAAGTCAACAGAAGCTGGCCTGAACGCC

|||||
TCAGAGAGAGGCTACGACCTGGCCCGCTGCGCCGTACAGGTCAAGAAAGCTGGGTGACAGCCG
X 10 20 30 40 50 60 70
GCCAGCCGATCTACCTGCTGACCTGTGACCGAGAAAGAGCTGTGCGCCACCATGAGTACCTGGTGGCCCTGC
|||
GCCAGCCGATCTACCTGCTGACCTGTGACCGAGAAAGAGCTGTGCGCCACCATGAGTACCTGGTGGCCCTGC
|||
GCCAGCCGATCTACCTGCTGACCTGTGACCGAGAAAGAGCTGTGCGCCACCATGAGTACCTGGTGGCCCTGC
80 90 100 110 120 130 140
ACGAGAGGCGCAGC-----CCACGATGACCGCTCCCGGCGGCTGAGAGTCCCGTGCAGGTGG
|||
ACGAGAGGCGCAGCAGGATGAGCCGCGCCCGCTGATGACTGTCCCGGCGGATGAGAGTGGCGGTGAGACCG
150 160 170 180 190 200 210
X
ACGACAT
|||||
ACGACAT
220 X

20. US-09-697-123b-19 (1-208)
US-09-697-123b-1 Sequence 13, Application US/09697123b

Initial Score = 143 Optimized Score = 169 Significance = -0.22
Residue Identity = 82% Matches = 183 Mismatches = 25
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAGAGGCTACGACCTGGCCCGCTGCGCCGTACAGGTCAAGAAAGCTGGGTGACAGCCG
|||
TCAGAGAGAGGCTACGACCTGGCCCGCTGCGCCGTACAGGTCAAGAAAGCTGGGTGACAGCCG
X 10 20 30 40 50 60 70
GCCAGCCGATCTACCTGCTGACCTGTGACCGAGAGAGCTGTGCGCCACCATGAGTACCTGGTGGCCCTGC
|||
ATCATTCGATTCACACACACACCTGTGACCGAGAGAGCTGTGCGCCACCATGAGTACCTGGTGGCCCTGC
80 90 100 110 120 130 140
AC-----GAGGCGCAGACAC--GATGACCGTCCCGGCGGCTGAGAGTCCCGTGGAGGTG
|||
ACCGAGGCTCTCAGGGGTGCGCAGCCCGCTTATGACTGTCCCGGCGGCTGAGAGTCCCGTGGAGAAACCG
150 160 170 180 190 200 210
X
ACGACAT
|||||
ACGACAT
220 X

21. US-09-697-123b-19 (1-208)
US-09-697-123b-1 Sequence 18, Application US/09697123b

Initial Score = 125 Optimized Score = 170 Significance = -0.59
Residue Identity = 83% Matches = 176 Mismatches = 32
Gaps = 3 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAGAGGCTACGACCTGGCCCGCTGCGCCGTACAGGTCAAGAAAGCTGGGTGAGAGCCG
|||
TCAGAGAGAGGCTACGACCTGGCCCGCTGCGCCGTACAGGTCAAGAAAGCTGGGTGAGAGCCG
X 10 20 30 40 50 60 70
GCCAGCCGATCTACCTGCTGACCTGTGACCGAGAGAGCTGTGCGCCACCATGAGTACCTGGTGGCCCTGC
|||
ACCGGCGCAGACGACCTGACAGCTGGCCCGCTGCGCCGTACAGGTCAAGAAAGCTGGGTGAGAGCCG
80 90 100 110 120 130 140
X
ACGACAT
|||||
ACGACAT
220 X

150 160 170 180 190 200
TGCAGAGAGGCGCAGACCGATGACCGTCCCGGCGGCTGAGAGTCCCGTGCAGGTGAGACCAT
|||
TGCATTCAGAGGCGCAGACCGATGACCGTCCCGGCTGAGAGTCCCGTGCAGGTGAGACCAT
150 160 170 180 190 200 210

22. US-09-697-123b-19 (1-208)
US-09-697-123b-2 Sequence 21, Application US/09697123b

Initial Score = 125 Optimized Score = 177 Significance = -0.59
Residue Identity = 85% Matches = 184 Mismatches = 24
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAGAGGCTACGACCTGGCCCGCTGCGCCGTACAGGTCAAGAAAGCTGGGTGAGAGCCG
|||
TCAGAGAGAGGCTACGACCTGGCCCGCTGCGCCGTACAGGTCAAGAAAGCTGGGTGAGAGCCG
X 10 20 30 40 50 60 70
GCCAGCCG-----GATCAGCTGCTGACCTGTGACCGAGAGAGCTGTGCGCCACCATGAGTACCTGGTGC
|||
CCAAATCGGCTCAGGTGACCCACACACCGCTTACCGAGAGAGCTGTGCGCCACCATGAGTACCTGGTGC
80 90 100 110 120 130 140

140 150 160 170 180 190 200
GCCGTGACAGAGGCGCAGACCATGATGACCGTCCCGGCGGCTGAGAGTCCCGTGGAGTGTGACAGCAT
|||
GCCGTGACAGAGGCGCAGACCATGATGACCGTCCCGGCGGCTGAGAGTCCCGTGGAGTGTGACAGCAT
150 160 170 180 190 200 210 X

23. US-09-697-123b-19 (1-208)
US-09-697-123b-1 Sequence 14, Application US/09697123b

Initial Score = 116 Optimized Score = 168 Significance = -0.78
Residue Identity = 81% Matches = 175 Mismatches = 33
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAGAGGCTACGACCTGGCCCGCTGCGCCGTACAGGTCAAGAAAGCTGGGTGAGAGCCG
|||
TCAGAGAGAGGCTACGACCTGGCCCGCTGCGCCGTACAGGTCAAGAAAGCTGGGTGAGAGCCG
X 10 20 30 40 50 60 70
GCCAGCCG-----ATCAGCTGCTGACCTGTGACCGAGAGAGCTGTGCGCCACCATGAGTACCTGGTGC
|||
CCGAGTGGCGCTGACCCGCTGCGCCGCTGACCGAGAGAGCTGTGCGCCACCATGAGTACCTGGTGC
80 90 100 110 120 130 140

140 150 160 170 180 190 200
GCCGTGACAGAGGCGCAGACCATGATGACCGTCCCGGCGGCTGAGAGTCCCGTGGAGTGTGAGCAT
|||
GCCGTGACAGAGGCGCAGACCATGATGACCGTCCCGGCGGCTGAGAGTCCCGTGGAGTGTGAGCAT
150 160 170 180 190 200 210 X

24. US-09-697-123b-19 (1-208)
US-09-697-123b-7 Sequence 7, Application US/09697123b

Initial Score = 115 Optimized Score = 177 Significance = -0.80
Residue Identity = 85% Matches = 184 Mismatches = 24
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAGAGGCTACGACCTGGCCCGCTGCGCCGTACAGGTCAAGAAAGCTGGGTGAGAGCCG
|||
TCAGAGAGAGGCTACGACCTGGCCCGCTGCGCCGTACAGGTCAAGAAAGCTGGGTGAGAGCCG
X 10 20 30 40 50 60 70
GCCAGCCG-----ATCAGCTGCTGACCTGTGACCGAGAGAGCTGTGCGCCACCATGAGTACCTGGTGC
|||
CCGAGTGGCGCTGACCCGCTGCGCCGCTGACCGAGAGAGCTGTGCGCCACCATGAGTACCTGGTGC
80 90 100 110 120 130 140

X
CGACAT
|||||
CGACAT
10 X

25. US-09-697-123B-19 (1-208)
US-09-697-123B-2 Sequence 25, Application US/09697123B

Initial Score	-	19	Optimized Score	-	19	Significance	-	-2.76
Residue Identity	=	100%	Matches	-	19	Mismatches	-	0
Gaps	=	0	Conservative Substitutions	-			-	0

X	10	X	
TCACGAGAGAGCGCTACGA			
TCACGAGAGAGCGCTACGA			
X	10	X	

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Query sequence being compared:	US-09-697-123B-18 (1-211)
Number of sequences searched:	26
Number of scores above cutoff:	25

Results of the initial comparison of US-09-697-123B-18 (1-211) with
File : US09697123B.seq

[illegible]

	Unary	K-tuple Joining penalty Window size
Similarity matrix	1	36
Mismatch penalty	5.00	20
Gap penalty	0.33	
Gap size penalty	12	
Cutoff score	0	
Randomization group		

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
---------	------	--------	--------------------

Times:	CPU	Total Elapsed
--------	-----	---------------

00:00:00:00

Number of residues: 5077

Number of scores above cutoff: 25

The scores below are sorted by initial score. Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

[illegible]

1. US-09-697-123B-1 Sequence 18, Application	211	211	2.90	0
--	-----	-----	------	---

Sequence Name	Description	Length	Score	Score	Sig.	Frame
---------------	-------------	--------	-------	-------	------	-------

[illegible]

1. US-09-697-123B-18 (1-211)
US-09-697-123B-1 Sequence 18, Application US/09697123B

Initial Score	-	211	Optimized Score	-	211	Significance	-	2.90
Residue Identity	-	100%	Matches	-	211	Mismatches	-	0
Gaps	-	0	Conservative Substitutions	-			-	0

TCGAAGGAGAAAGCGCTACGACCTGGCCCGCTGGGTGGGTACAAAGTCAACAAAGCTGGGATCACCAGAG
X 10 20 30 40 50 60 70
TCGAAGGAGAAAGCGCTACGACCTGGCCCGCTGGGTGGGTACAAAGTCAACAAAGTGGGATCACCAGAG
X 10 20 30 40 50 60 70

ACCGGCCGACGACGCTTCGACGACGACCGAAGAGACGCTGTCGCCACCATGAGTACTGTTGGG
 80 90 100 110 120 130 140
 ACCGGCCGACGACGCTTCGACGACGACCGAAGAGACGCTGTCGCCACCATGAGTACTGTTGGG
 80 90 100 110 120 130 140
 ACCGGCCGACGACGCTTCGACGACGACCGAAGAGACGCTGTCGCCACCATGAGTACTGTTGGG
 80 90 100 110 120 130 140

150	160	170	180	190	200	210
TCATCAGGGCGCAAGACGATGACCGTCCGGGTGGAGTGCAGGTCGACGACATAT						
150	160	170	180	190	200	210
TGATCAGGGCGCAAGACGATGACCGTCCGGGTGGAGTGCAGGTCGACGACATAT						

2. US-09-697-123B-18 (1-211)
US-09-697-123B-1 Sequence 19, Application US/09697123B

Initial score	-	125	Optimized score	-	170	Significance	-	0.43
Residue Identity	-	83%	Matches	-	176	Mismatches	-	32
Gaps	-	3	Conservative Substitutions	-			-	0

[illegible]

8. US-09-697-123B-18 (1-211)
US-09-697-123B-8 Sequence 8, Application US/09697123B

Initial Score =	117	Optimized Score =	164	Significance =	0.20
Residue Identity =	80%	Matches =	170	Mismatches =	38
Gaps =	3	Conservative Substitutions			0

X
10
20
30
40
50
60
70
 TCAAGAGAGACCCATCGACCTGCGGCCCGCTGGGTGATCAAGGTCAACAGAGCTGGGGCTATCCACGAG
 TCAAGAGAGACCCATCGACCTGCGGCCCGCTGGGTGATCAAGGTCAACAGAGCTGGGGCTATCCACGAG
 X
10
20
30
40
50
60
70

ACCGGGGCAGACAGCTTCGACACACCGTCAGAGGAGAGCTGCTCCGACACATCAGATACCTGGTGGGG
 80 90 100 110 120 130 140
 - - - - -
 ATCATTC---GATCACACAGAGACGTCAGACCCAGAAAGACGCTGCCACCATCAGATATGGTCCGGCC
 80 90 100 110 120 130 140

TGCATTCAGGGCGGCAAGACGATGACCGCTCCCGGCTGAGATCGAGGTGCGCTTCGACACACT
 150 160 170 180 190 200 210
 TGCACGGGGCCAGGCCACGACGATGACCTGGCCGGGGCGGGGTTCAGAGTGGCCGGTGAACCCAGACACT
 150 160 170 180 190 200 X

9. US-09-697-123B-18 (1-211)
US-09-697-123B-2 Sequence 2, Application US/09697123B

Initial Score	=	117	Optimized Score	=	161	Significance	=	0.20
Residue Identity	=	79%	Matches	=	167	Mismatches	=	41
Gaps	=	3	Conservative Substitutions	=			=	0

X
 10
 20
 30
 40
 50
 60
 70
 TCAAGCAAGCGCTACACATGGCCCGCTGGTGGTCAAGTCAACGAAGTGGGATCACCAGAA
 TCAAGGAAAGCGCTACACCTGGCCGGGTGGCGGTACACAGTCAACGAGAGCTGGTCTAAGCTGC
 10 20 30 40 50 60 70

80
 90
 100
 110
 120
 130
 140
 ACCGGCCGACACGACGCTGACCCACCGAGGAGGAGCGTCGTCCAGCAATGATGACCTGTTGCGGC
 GCGAAGCGGATCCACGACGCTGCG---ACGCTGACCGAGGAGACGCTGTATGACCAACCATGATGACTGTTGCGGC
 80 90 100 110 120 130 140

TCGATCAGGGGCACAAACGATGACCCCTCCGGGTGGATCGAGTCCCGTCGAGGTGCACACAT
 |||| | ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 TGCACAGGGGTAGTCGGCCGTACGGTTCCGGGGGCCGAGGTCCCGTCGAGACACGACACAT
 150 160 170 180 190 200 X

10. US-09-697-123B-18 (1-211)
US-09-697-123B-1 Sequence 13, Application US/09697123B

Initial Score =	116	Optimized Score =	152	Significance =	0.17
Residue Identity =	73%	Matches =	167	Mismatches =	41
Gaps =	18	Conservative Substitutions			0

TC AAG AAG AAG CGC TAC CAG TGG CCG CCG GCG TGG CCG TAC AAG GTA CCA AAG AAG CCG GCG GAT CAC CAG A
10 20 30 40 50 60 70
TC AAG GAG AAG CGC TAC CAG TGG CCG CCG GCG TGG CCG TAC AAG GTA CCA AAG AAG CCG GCG GAT CAC CAG A
10 20 30 40 50 60 70

ACCCGCGACACGACGCTCGACCCAGCAGAGAGAGCTGTGTGGCCACATCGAGTACTGTGTGGGC
 80 90 100 110 120 130 140
 ATCATTCATTCACACACC---ACAGCCTACCGAAGAGAGCTGTGTGGCCACATCGAGTACTGTGTGGGC
 80 90 100 110 120 130 140

[illegible]

210
TCGACGACAT
| | | | | | | |
CCGACGACAT
220 >

11. US-09-697-123B-18 (1-211)
US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score	=	116	Optimized Score	=	150	Significance	=	0.17
Residue Identity	=	73%	Matches	=	165	Mismatches	=	43
Gaps	=	18	Conservative Substitutions	=			=	0

TCAAGGAGAGCCCTACGACTCGGCCGCCTGGTGTGAAGAATCAAGAAGTCGGGSCATCTACCGAG
|||||
|||
TCAGGAGAGCCCTACGACTCGGCCGCCTGTAAAGGTCAACAAGAACCTGGTGCTGCACCCGCC
X 10 20 30 40 50 60 70

ACCCGGCGACGACGTCTGCACCAAGGAGACCTCTCCACATTCACAGTCTGGTGGGG
 60 90 100 110 120 130 140
 --- 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 GCGAGCCGATCAG---TGTGTCCAGCGTGCAGGAGACCGTCTCCGCGACATTCGAATACCTGTCGGG
 80 90 100 110 120 130 140

150 160 170 180 190 200 210
 TGCATCAGGCGCAG-----ACGATGACCTCCCGGGTGGATCGAGTCCCGCTCGAGT
 ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 TGCACACGCGCGTACGATGCGCAGCCCGCGCTCATGTACTCTCCCGCGCATCGAGTGGCGGTGACAC

TCGACGACAT
| | | | |
CCGACGACAT
220 X

US-09-697-123B-1 Sequence 14, Application US/09697123B

Initial score	=	115	Optimized score	=	166	Significance	=	0.14
Residue Identity	=	80%	Matches	=	172	Mismatches	=	39
Gaps	=	3	Conservative Substitutions	=			=	0

TC AAGGAGAGCGCTACGACCTGCGCCGCTGCTGGGTACAAAGTTCAAACAAGCTGGGCTACCGAG-
 |||||
 TC AAGGAGAGCGCTACGACCTGCGCCGCTGCTGGGTACAAAGTTCAAACAAGCTGGGCTACCGAG-
 |||||
 TC AAGGAGAGCGCTACGACCTGCGCCGCTGCTGGGTACAAAGTTCAAACAAGCTGGGCTACCGAG-
 |||||

-7AAGCGGCGCGACGACGACCTCGACACCGGTGACCGAAGAGAGCGTGTGCCACCATCATCTGTCGTGC
 CGAGTGTGGGCCCTACCCGCCCTCGACGACGCGTGCACGAAGCGAGTGTGTCGCCCATGTGATACCTGTGTC
 80 90 100 110 120 130 140

[illegible]

13. US-09-697-123B-18 (1-211)
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial score	-	114	Optimized Score	-	158	Significance	-	0.11
Residue Identity	-	77%	Matches	-	164	Mismatches	-	44
Gaps	-	3	Conservative Substitutions	-			-	0

X
10 20 30 40 50 60 70
TCAAGAGAACCGCTACGACCTGGCCCGCGGTGCGTACAGCTTAACAAGAGCTGGGCAATCACCCGAGA
|||||
TCAAGAGAACCGCTACGACCTGGCCCGCGGTGCGTAAAGCTTAACAAGAGCTGGGCAATCACCCGAGA
X
10 20 30 40 50 60 70
ACCCGGCCGACAGCAGCTGCGACCGCTGACCGGAGAGAGAGCTGCTGCGCCACCATGAGTACCTGGTGGGC
|||
GCGAGCCATCAGC---TCGTCGACGCTGACCGGAGAGAGAGAGCTGCTGCGCCACCATGAGTACCTGGTGGGC
80 90 100 110 120 130 140
TGCATCAGGCGGCAAGAGAGATGACCGTCCCGGGTGGAGTGCAGGTGCGCCCGCTGAGAGCTGACAGCAT
|||||
TGCAGAGGCTCAGACCACTATGATCTTCCGGGCGCGCTGAGGTGCGCGTGAAGAACCGACGACAT
X
150 160 170 180 190 200
TGCATCAGGCGGCAAGAGAGATGACCGTCCCGGGTGGAGTGCAGGTGCGCCCGCTGAGAGCTGACAGCAT
|||||
TGCAGAGGCTCAGACCACTATGATCTTCCGGGCGCGCTGAGGTGCGCGTGAAGAACCGACGACAT
X
150 160 170 180 190 200

14. US-09-697-123B-18 (1-211)

US-09-697-123B-1 Sequence 1, Application US/09697123B

Initial Score = 114 Optimized Score = 162 Significance = 0.11
Residue Identity = 79% Matches = 168 Mismatches = 40
Gaps = 3 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAACCGCTACGACCTGGCCCGCGGTGCGTACAGCTTAACAAGAGCTGGGCAATCACCCGAGA
|||||
TCAAGAGAACCGCTACGACCTGGCCCGCGGTGCGTACAGCTTAACAAGAGCTGGGCAATCACCCGAGA
X
10 20 30 40 50 60 70
ACCCGGCCGACAGCAGCTGCGACCGCTGACCGGAGAGAGAGCTGCTGCGCCACCATGAGTACCTGGTGGGC
|||
GCGATCGCATCAGCAGCT---CCAGCTGACCGAGAGAGAGAGCTGCTGCGCCACCATGAGTACCTGGTGGGC
80 90 100 110 120 130 140
TGCATCAGGCGGCAAGAGAGATGACCGTCCCGGGTGGAGTGCAGGTGCGCCCGCTGAGAGCTGACAGCAT
|||||
TGCAGAGGCTCAGACCACTATGATCTTCCGGGCGCGCTGAGGTGCGCGTGAAGAACCGACGACAT
X
150 160 170 180 190 200

150 160 170 180 190 200 210
TGCATCAGGCGGCAAGAGAGATGACCGTCCCGGGTGGAGTGCAGGTGCGCCCGCTGAGAGCTGACAGCAT
|||||
TGCAGAGGCTCAGACCACTATGATCTTCCGGGCGCGCTGAGGTGCGCGTGAAGAACCGACGACAT
X
150 160 170 180 190 200

15. US-09-697-123B-18 (1-211)

US-09-697-123B-2 Sequence 23, Application US/09697123B

Initial Score = 113 Optimized Score = 165 Significance = 0.09
Residue Identity = 81% Matches = 171 Mismatches = 37
Gaps = 3 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAACCGCTACGACCTGGCCCGCGGTGCGTACAGCTTAACAAGAGCTGGGCAATCACCCGAGA
|||||
TCAAGAGAACCGCTACGACCTGGCCCGCGGTGCGTACAGCTTAACAAGAGCTGGGCAATCACCCGAGA
X
10 20 30 40 50 60 70
ACCCGGCCGACAGCAGCTGCGACCGCTGACCGGAGAGAGAGCTGCTGCGCCACCATGAGTACCTGGTGGGC
|||
GTCAGGCGGATCAGCAGCTCG---ACGCTGACCGAGAGAGAGAGCTGCTGCGCCACCATGAGTACCTGGTGGGC
80 90 100 110 120 130 140
TGCATCAGGCGGCAAGAGAGATGACCGTCCCGGGTGGAGTGCAGGTGCGCCCGCTGAGAGCTGACAGCAT
|||||
TGCAGAGGCTCAGACCACTATGATCTTCCGGGCGCGCTGAGGTGCGCGTGAAGAACCGACGACAT
X
150 160 170 180 190 200

Initial Score = 113 Optimized Score = 169 Significance = 0.09
Residue Identity = 82% Matches = 175 Mismatches = 33
Gaps = 3 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAACCGCTACGACCTGGCCCGCGGTGCGTACAGCTTAACAAGAGCTGGGCAATCACCCGAGA
|||||
TCAAGAGAACCGCTACGACCTGGCCCGCGGTGCGTACAGCTTAACAAGAGCTGGGCAATCACCCGAGA
X
10 20 30 40 50 60 70
ACCCGGCCGACAGCAGCTTGCACCGCTGACCGGAGAGAGAGCTGCTGCGCCACCATGAGTACCTGGTGGGC
|||
CCGCGTCCCGATCAGACAGCAGCACTGACCGGAGAGAGAGAGCTGCTGCGCCACCATGAGTACCTGGTGGGC
80 90 100 110 120 130 140
TGCATCAGGCGGCAAGAGAGATGACCGTCCCGGGTGGAGTGCAGGTGCGCCCGCTGAGAGCTGACAGCAT
|||||
TGCAGAGGCTCAGACCACTATGATCTTCCGGGCGCGCTGAGGTGCGCGTGAAGAACCGACGACAT
X
150 160 170 180 190 200

150 160 170 180 190 200 210
TGCATCAGGCGGCAAGAGAGATGACCGTCCCGGGTGGAGTGCAGGTGCGCCCGCTGAGAGCTGACAGCAT
|||||
TGCAGAGGCTCAGACCACTATGATCTTCCGGGCGCGCTGAGGTGCGCGTGAAGAACCGACGACAT
X
150 160 170 180 190 200

17. US-09-697-123B-18 (1-211)

US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 113 Optimized Score = 159 Significance = 0.09
Residue Identity = 78% Matches = 165 Mismatches = 43
Gaps = 3 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAACCGCTACGACCTGGCCCGCGGTGCGTACAGCTTAACAAGAGCTGGGCAATCACCCGAGA
|||||
TCAAGAGAACCGCTACGACCTGGCCCGCGGTGCGTACAGCTTAACAAGAGCTGGGCAATCACCCGAGA
X
10 20 30 40 50 60 70
ACCCGGCCGACAGCAGCTTGCACCGCTGACCGGAGAGAGAGCTGCTGCGCCACCATGAGTACCTGGTGGGC
|||
GCGATCGCATCAGCAGCT---TCGTCGACGCTGACCGGAGAGAGAGCTGCTGCGCCACCATGAGTACCTGGTGGGC
80 90 100 110 120 130 140
TGCATCAGGCGGCAAGAGAGATGACCGTCCCGGGTGGAGTGCAGGTGCGCCCGCTGAGAGCTGACAGCAT
|||||
TGCAGAGGCTCAGACCACTATGATCTTCCGGGCGCGCTGAGGTGCGCGTGAAGAACCGACGACAT
X
150 160 170 180 190 200

150 160 170 180 190 200 210
TGCATCAGGCGGCAAGAGAGATGACCGTCCCGGGTGGAGTGCAGGTGCGCCCGCTGAGAGCTGACAGCAT
|||||
TGCAGAGGCTCAGACCACTATGATCTTCCGGGCGCGCTGAGGTGCGCGTGAAGAACCGACGACAT
X
150 160 170 180 190 200

18. US-09-697-123B-18 (1-211)

US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 113 Optimized Score = 159 Significance = 0.09
Residue Identity = 78% Matches = 165 Mismatches = 43
Gaps = 3 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAACCGCTACGACCTGGCCCGCGGTGCGTACAGCTTAACAAGAGCTGGGCAATCACCCGAGA
|||||
TCAAGAGAACCGCTACGACCTGGCCCGCGGTGCGTACAGCTTAACAAGAGCTGGGCAATCACCCGAGA
X
10 20 30 40 50 60 70
ACCCGGCCGACAGCAGCTTGCACCGCTGACCGGAGAGAGAGCTGCTGCGCCACCATGAGTACCTGGTGGGC
|||
GCGAGCCATCAGCAGCT---TCGTCGACGCTGACCGGAGAGAGAGCTGCTGCGCCACCATGAGTACCTGGTGGGC
80 90 100 110 120 130 140
TGCATCAGGCGGCAAGAGAGATGACCGTCCCGGGTGGAGTGCAGGTGCGCCCGCTGAGAGCTGACAGCAT
|||||
TGCAGAGGCTCAGACCACTATGATCTTCCGGGCGCGCTGAGGTGCGCGTGAAGAACCGACGACAT
X
150 160 170 180 190 200 210

16. US-09-697-123B-18 (1-211)

US-09-697-123B-1 Sequence 17, Application US/09697123B

[illegible]

A 10 20 30 40 50 60 70
 TTTAAGGAAACCGCTACGACTGCGCCGGTGGTGGTCAAGTCAAGAAGCTGGGCTATCCGAGAG
 TTTAAGGAAACCGCTACGACTGG-TGGGTGGCGCCGTTACAAAGTCACAAAAAGTCCGTGAACCTGCG
 X 10 20 30 40 50 60 70

X 10 20 30 40 50 60 70
 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGGTACAAAGTCAACAAGAGCTCGCGCTGAACACCG
 TCAAGGAGAGCGCTACGACCTGGCGCGGTGGCGGTACAAAGTCAACAAGAGCTCGCGCTGAACACCG
 X 10 20 30 40 50 60 70

```
80      90      100      110      120      130      140
CGTCCCGGATCACAGACACCTCTGACCGAAGAGGACGTGTCGCCACCATCGAGTACCTGGTCCGCTGC
|||||
ATCATCGGATCACACACACGACGCTTACCCGAAGAAGACGTCGTGCCACCATCGAGTATCTGGTCCGCTGC
80      90      100      110      120      130      140
150      160      170      180      190      200      210
ACGAGGGCCACACACCATCACCTCCCGGGCGAGTCGAGGTGCCGGTGGAAACCGAGACAT
|||||
ACGAGGGCCAGGCGCACCATGACCTGCGGGCGGGTTCGAGGTGCCGGTGGAAACCGAGACAT
150      160      170      180      190      200      210
X
```

3. US-09-697-123B-17 (1-208)

US-09-697-123B-1 Sequence 19, Application US/09697123B

Initial Score = 183 Optimized Score = 183 Significance = 0.64
Residue Identity = 87% Matches = 183 Mismatches = 25
Gaps = 0 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCGGCCCTGAACACCG
|||||
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCGGCCCTGAACACCG
X      10      20      30      40      50      60      70
```

```
80      90      100      110      120      130      140
CGTCCCGGATCACAGACACCTCTCACCGAAGAGGACGTGCTGCCACCATCGAGTACCTGGTCCGCTGC
|||||
GCCAGCCGATCACGTGCTGACCTCTACCGAGGAAGACGTCGTGCCACCATCGAGTACCTGGTCCGCTGC
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200      210
ACGAGGGCCACACACGATGACCTCCCGGGCGAGTCGAGGTGCCGGTGGAAACCGAGACAT
|||||
ACGAGGGCCAGACGATGACCTCCCGGGCGGTCGAGGTGCCGGTGGAAACCGAGACAT
150      160      170      180      190      200      210
X
```

4. US-09-697-123B-17 (1-208)

US-09-697-123B-1 Sequence 15, Application US/09697123B

Initial Score = 183 Optimized Score = 183 Significance = 0.64
Residue Identity = 87% Matches = 183 Mismatches = 25
Gaps = 0 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCGGCCCTGAACACCG
|||||
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCGGCCCTGAACACCG
X      10      20      30      40      50      60      70
```

```
80      90      100      110      120      130      140
CGTCCCGGATCACAGACACCTCTGACCGAAGAGGACGTGTCGCCACCATCGAGTACCTGGTCCGCTGC
|||||
GTGAGCGGATCACGAGCTGACCGTACCGAGGAGACGTCGTGCCACCATCGAGTACCTGGTCCGCTGC
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200      210
ACGAGGGCCACACGATGACCTCCCGGGCGAGTCGAGGTGCCGGTGGAAACCGAGACAT
|||||
ACGAGGGTCAGGCCACGATGACCTCCCGGGCGGTCGAGGTGCCGGTGGAAACCGAGACAT
150      160      170      180      190      200      210
X
```

5. US-09-697-123B-17 (1-208)

US-09-697-123B-1 Sequence 1, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.62
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
```

```
TCAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCGGCCCTGAACACCG
|||||
TCAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCGGCCCTGAACACCG
X      10      20      30      40      50      60      70
```

```
80      90      100      110      120      130      140
CGTCCCGGATCACAGACACCTCTGACCGAAGAGGACGTGTCGCCACCATCGAGTACCTGGTCCGCTGC
|||||
GCGATCCGATCACACGCTCCACGCTGACCGGAGGACGTCGTGCCACCATCGAGTACCTGGTCCGCTGC
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200      210
ACGAGGGCCACACACGATGACCTCCCGGGCGAGTCGAGGTGCCGGTGGAAACCGAGACAT
|||||
ACGAGGGCCAGCAGACGATGACCTCCCGGGCGGACCGAGGTGCCGGTGGAGACCGAGACAT
150      160      170      180      190      200      210
X
```

6. US-09-697-123B-17 (1-208)

US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.60
Residue Identity = 87% Matches = 181 Mismatches = 27
Gaps = 0 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCGGCCCTGAACACCG
|||||
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCGGCCCTGAACACCG
X      10      20      30      40      50      60      70
```

```
80      90      100      110      120      130      140
CGTCCCGGATCACAGACACCTCTGACCGAAGAGGACGTGTCGCCACCATCGAGTACCTGGTCCGCTGC
|||||
AGAATCGGCAACCCACGACCTGACCGAGAGGACGTCGTGCCACCATCGAATACCTGGTCCGCTGC
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200      210
ACGAGGGCCACACGATGACCTCCCGGGCGAGTCGAGGTGCCGGTGGAAACCGAGACAT
|||||
ACGAGGGCCAGCAGACGATGACCTCCCGGGTGGCGGTCGAGGTGCCGGTGGAGACCGAGACAT
150      160      170      180      190      200      210
X
```

7. US-09-697-123B-17 (1-208)

US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.60
Residue Identity = 87% Matches = 181 Mismatches = 27
Gaps = 0 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCGGCCCTGAACACCG
|||||
TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCGGCCCTGAACACCG
X      10      20      30      40      50      60      70
```

```
80      90      100      110      120      130      140
CGTCCCGGATCACAGACACCTCTGACCGAAGAGGACGTGTCGCCACCATCGAGTACCTGGTCCGCTGC
|||||
GCCAGCCGATCACGAGCTCGAGCTGACCGAGGAGACGTCGTGCCACCATCGAATACCTGGTCCGCTGC
80      90      100      110      120      130      140
```

```
150      160      170      180      190      200      210
ACGAGGGCCACACGATGACCTCCCGGGCGAGTCGAGGTGCCGGTGGAAACCGAGACAT
|||||
ACGAGGGCCAGCAGGATGACCTCCCGGGCGGTCGAGGTGCCGGTGGAGACCGAGACAT
150      160      170      180      190      200      210
X
```

8. US-09-697-123B-17 (1-208)

US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 179 Optimized Score = 179 Significance = 0.56
Residue Identity = 86% Matches = 179 Mismatches = 29
Gaps = 0 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGTGGCGCTGAACACCG
|||||
TCAAGGAGAAGCGCTACGACCTCGCGCTGGTGGTCAAGAGTCAACAAGAGTGGCGCTGAACACCG
X 10 20 30 40 50 60 70
```

```
80 90 100 110 120 130 140
CGTCCCGGATCAGACACGACCTCTGACCGAAGAGAGAGTGGCGGTGGAACCGAGAGACAT
|||||
GCCAGCCCATCACCAGCTCGACGCTGACCGAGAGAGAGTGGCGGACCATCGAATACCTGGTCCGCTTGC
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200 X
ACGAGGCCACACGATGACCTCGCGGGGAGTTCGAGGTGGCGGTGGAACCGAGAGACAT
|||||
ACGAGGCCACGATGACCTCGCGGGGAGTTCGAGGTGGCGGTGGAACCGAGAGACAT
150 160 170 180 190 200 X
```

9. US-09-697-123B-17 (1-208)

US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 179 Optimized Score = 179 Significance = 0.56
Residue Identity = 86% Matches = 179 Mismatches = 29
Gaps = 0 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGTGGCGCTGAACACCG
|||||
TCAAGGAGAAGCGCTACGACCTCGCGGTGGTGGTCAAGAGTCAACAAGAGTGGCGCTGAACACCG
X 10 20 30 40 50 60 70
```

```
80 90 100 110 120 130 140
CGTCCCGGATCAGACACGACCTCTGACCGAAGAGAGAGTGGCGGTGGAACCGAGAGACAT
|||||
GCGATCCGATCACCAGCTCCACGCTGACCGAGAGAGAGTGGCGGACCATCGAATACCTGGTCCGCTTGC
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200 X
ACGAGGCCACACGATGACCTCGCGGGGAGTTCGAGGTGGCGGTGGAACCGAGAGACAT
|||||
ACGAGGTTCAGACACGATGACCTCGCGGGGAGTTCGAGGTGGCGGTGGAACCGAGAGACAT
150 160 170 180 190 200 X
```

10. US-09-697-123B-17 (1-208)

US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 178 Optimized Score = 178 Significance = 0.53
Residue Identity = 85% Matches = 178 Mismatches = 30
Gaps = 0 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGTGGCGCTGAACACCG
|||||
TCAAGGAGAAGCGCTACGACCTCGCGGTGGTGGTCAAGAGTCAACAAGAGTGGCGCTGAACACCG
X 10 20 30 40 50 60 70
```

```
80 90 100 110 120 130 140
CGTCCCGGATCAGACACGACCTCTGACCGAAGAGAGAGTGGCGGTGGAACCGAGAGACAT
|||||
GCGAGCCCATCAGCTCGACGCTGACCGAAGAGAGAGTGGCGGACCATCGAATATCTGGTCCGCTTGC
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200 X
ACGAGGCCACACGATGACCTCGCGGGGAGTTCGAGGTGGCGGTGGAACCGAGAGACAT
|||||
ACGAGGTTCAGACACGATGACCTCGCGGGGAGTTCGAGGTGGCGGTGGAACCGAGAGACAT
150 160 170 180 190 200 X
```

11. US-09-697-123B-17 (1-208)

US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 178 Optimized Score = 178 Significance = 0.53
Residue Identity = 85% Matches = 178 Mismatches = 30
Gaps = 0 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGTGGCGCTGAACACCG
|||||
TCAAGGAGAAGCGCTACGACCTCGCGGTGGTGGTCAAGAGTCAACAAGAGTGGCGCTGAACACCG
X 10 20 30 40 50 60 70
```

```
80 90 100 110 120 130 140
CGTCCCGGATCAGACACGACCTCTGACCGAAGAGAGAGTGGCGGTGGAACCGAGAGACAT
|||||
GCGAGCCCATCAGCTCGTGGAGCTGACCGTGGCGGTGGCGGTGGAACCGAGAGACAT
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200 X
ACGAGGCCACACGATGACCTCGCGGGGAGTTCGAGGTGGCGGTGGAACCGAGAGACAT
|||||
ACGAGGTTCAGACACGATGACCTCGCGGGGAGTTCGAGGTGGCGGTGGAACCGAGAGACAT
150 160 170 180 190 200 X
```

12. US-09-697-123B-17 (1-208)

US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 178 Optimized Score = 178 Significance = 0.53
Residue Identity = 85% Matches = 178 Mismatches = 30
Gaps = 0 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGGTACAAGGTCAACAAGAGTGGCGCTGAACACCG
|||||
TCAAGGAGAAGCGCTACGACCTCGCGGTGGTGGTCAAGAGTCAACAAGAGTGGCGCTGAACACCG
X 10 20 30 40 50 60 70
```

```
80 90 100 110 120 130 140
CGTCCCGGATCAGACACGACCTCTGACCGAAGAGAGAGTGGCGGTGGAACCGAGAGACAT
|||||
GCGAGCCCATCAGCTCGTGGAGCTGACCGTGGCGGTGGCGGTGGAACCGAGAGACAT
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200 X
ACGAGGCCACACGATGACCTCGCGGGGAGTTCGAGGTGGCGGTGGAACCGAGAGACAT
|||||
ACGAGGTTCAGACACGATGACCTCGCGGGGAGTTCGAGGTGGCGGTGGAACCGAGAGACAT
150 160 170 180 190 200 X
```

13. US-09-697-123B-17 (1-208)

US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 177 Optimized Score = 177 Significance = 0.51
Residue Identity = 85% Matches = 177 Mismatches = 31
Gaps = 0 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGGTACAAGGTCAACAAGAGTGGCGCTGAACACCG
|||||
TCAAGGAGAAGCGCTACGACCTCGCGGTGGTGGTCAAGAGTCAACAAGAGTGGCGCTGAACACCG
X 10 20 30 40 50 60 70
```

```
80 90 100 110 120 130 140
CGTCCCGGATCAGACACGACCTCTGACCGAAGAGAGAGTGGCGGTGGAACCGAGAGACAT
|||||
GCGAGCCCATCAGCTCGTGGAGCTGACCGAAGAGAGAGTGGCGGACCATCGAATATCTGGTCCGCTTGC
80 90 100 110 120 130 140
```



```
|||||
TCAAGGAGAGCGCTACGACCTGGCCCGCTCGCGCTACAAAGGTCAACAAGAGCTGGGCTGAACACCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CGTCCCGGATCAGCAGGACCACTCTGACCGAAGAGACGTGCTCGCCACCATCGAGTACTGTCGGCTGCG
|||||
ATCATCGGATCACCAACGACGCTCACCGAAGAGACGTGCTCGCCACCATCGAGTACTGTCGGCTGCG
80 90 100 110 120 130 140

150 160 170 180 190 200
ACGAGGCG-----CACACCACGATGACCGTCCCGCGGAGTCGAGGTCCCGGTGGAACCG
|||||
ACACGCGCTCTCAGGTGGCGAGCGCCCGTTATGACTGTCCCGCGGGTCTGAGGTCCCGGTGGAACCG
150 160 170 180 190 200 210

X
ACGACAT
ACGACAT
220 X
```

20. US-09-697-123B-17 (1-208)

US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score = 137 Optimized Score = 165 Significance = -0.31
Residue Identity = 80% Matches = 172 Mismatches = 36
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGCTCGGCT-----GA
|||||
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGGGTGGCGGTACAAGGTGAACAAGAGCTGGTCTTGGCGGTG
X 10 20 30 40 50 60 70
```

```
70 80 90 100 110 120 130
ACACCGGTCCTCCGATCACGACCACTCTGACCGAAGAGACGTGCTCGCCACCATCGAGTACTGCTGCC
|||||
CCAACCGGCTCTGTGACTGCTGACCTGACCGAGGAGACGTGCTCGCCACCATCGGCTACTGCTGCTGC
80 90 100 110 120 130 140
```

```
140 150 160 170 180 190 200 X
GCCTGCACGAGGGCCACACGATGACCTGCTCCCGGGAGTCGAGGTCCCGGTGGAACCGACACAT
|||||
GCCTGCACGAGGGCCACACGATGACCTGACCGCGGCTCGAGGTCCCGGTGAGGTGCGACACAT
150 160 170 180 190 200 210 X
```

21. US-09-697-123B-17 (1-208)

US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 137 Optimized Score = 164 Significance = -0.31
Residue Identity = 78% Matches = 174 Mismatches = 34
Gaps = 15 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGCTCGGCTGAACACCG
|||||
TCAAGGAGAAGCGCTACGACCTGCCCCGCTCGCGCTACAAGGTCAACAAGAGCTGGGTCTGACGCGCG
X 10 20 30 40 50 60 70
```

```
80 90 100 110 120 130 140
CGTCCCGGATCACGACCACTCTGACCGAAGAGACGTGCTCGCCACCATCGAGTACTGCTCGGCTGCG
|||||
GCGAGCGGATCACGCTGACCGTGGCCCGCTCGCGCTACAAGGTCAACAAGAGCTGGGTCTGACGCGCG
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200
ACGAGGCG-----CACACCACGATGACCGTCCCGGGGAGTCGAGGTCCCGGTGGAACCG
|||||
ACACGCGCTCTCAGGTGGCGAGCGCCCGTTATGACTGTCTCCCGCGGGTCTGAGGTCCCGGTGGAACCG
150 160 170 180 190 200 210
```

```
X
ACGACAT
|||||
ACGACAT
220 X
```

22. US-09-697-123B-17 (1-208)

US-09-697-123B-1 Sequence 18, Application US/09697123B

Initial Score = 113 Optimized Score = 169 Significance = -0.80
Residue Identity = 82% Matches = 175 Mismatches = 33
Gaps = 3 Conservative Substitutions = 0

```
X 10 20 30 40 50 60
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGCTACAAGGTCAACAAGAGCTCGGCTGA-----ACA
|||||
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGGTGCGGTACAAGGTCAACAAGAGCTGGGCTACCGGAGA
X 10 20 30 40 50 60 70
```

```
70 80 90 100 110 120 130 140
CCGCTCCCGGATCACGACCACTTGACCGAAGAGACGTGCTGCCACCATCGAGTACTGTCGGTCCGCC
|||||
ACCCGCGCGACACGACCTCGACCGCTGACCGAGAGAGAGGTGCTGTCGCCACCATCGAGTACTGTCGGC
80 90 100 110 120 130 140
```

```
150 160 170 180 190 200 X
TCCACGAGGGCCACACGATGACCGTCCCGGGGAGTCGAGGTCCGGTGGAAACCGACGACAT
|||||
TGCATCAGGGCGCACAGACGATGACCGTCCCGGGTGGAGTGGAGTCCCGTTCGAGGTTCGACGACAT
150 160 170 180 190 200 210
```

23. US-09-697-123B-17 (1-208)

US-09-697-123B-2 Sequence 21, Application US/09697123B

Initial Score = 111 Optimized Score = 169 Significance = -0.84
Residue Identity = 82% Matches = 176 Mismatches = 32
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGGTGCGCTACAAGGTCAACAAGAGCTCGGCTG-----A
|||||
TCAAGGAGAAGCGCTACGATCTGCGCCCGTGGTTCGCTACAAGGTGAACAAGAGCTGGCGTGGCGGCA
X 10 20 30 40 50 60 70
```

```
70 80 90 100 110 120 130
ACACCGGTCCTCCGATCACGACCACTTGACCGAAGAGACGTGCTGCCACCATCGAGTACTGCTGCC
|||||
CCAAATCCGGCTCAGGTGACCAACCACTTCAACGAGGAGACGTGCTGCCACCATCGAGTACTGCTGCC
80 90 100 110 120 130 140
```

```
140 150 160 170 180 190 200 X
GCCTGCACGAGGGCCACACGATGACCGTCCCGGGGAGTCGAGGTGCGGTGGAAACCGACGACAT
|||||
GCCTGCACGAGGGCCACACGATGACCGTCCCGGGGAGTCGAGGTGCGGTGGAGTGTGAGACGACAT
150 160 170 180 190 200 210 X
```

24. US-09-697-123B-17 (1-208)

US-09-697-123B-1 Sequence 14, Application US/09697123B

Initial Score = 109 Optimized Score = 169 Significance = -0.88
Residue Identity = 82% Matches = 176 Mismatches = 32
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGGTGCGGTACAAGGTCAACAAGAGCTCGGCTGAACACCG
|||||
TCAAGGAGAAGCGCTACGACCTGCGCCAGGTTGGCGGTGCGGTACAAGGTCAACAAGAGCTCGGCTGCCGCGG
X 10 20 30 40 50 60 70
```


	10	20	30	40	50	60	70
X							
80		90	100	110	120	130	140
GCAGAGCCCATCAGCTCTCGACGCTGACCCGAAGAAGCGTCGGGACACATGAAATATCTGCTCGGTGG							
GCAGAGCCCATCAGCTCTCGACGCTGACCCGAAGAAGCGTCGGGACACATGAAATATCTGCTCGGTGG							
GCAGAGCCCATCAGCTCTCGACGCTGACCCGAAGAAGCGTCGGGACACATGAAATATCTGCTCGGTGG							
80		90	100	110	120	130	140
GCAGAGCCCATCAGCTCTCGACGCTGACCCGAAGAAGCGTCGGGACACATGAAATATCTGCTCGGTGG							
GCAGAGCCCATCAGCTCTCGACGCTGACCCGAAGAAGCGTCGGGACACATGAAATATCTGCTCGGTGG							
GCAGAGCCCATCAGCTCTCGACGCTGACCCGAAGAAGCGTCGGGACACATGAAATATCTGCTCGGTGG							
150	160	170	180	190	200		
ACGAGAGCTCAGACACCGATGACGCTTCGGGGGCGCGTCGAGTCCGGGGAACCGAGACGACAT							
ACGAGAGCTCAGACACCGATGACGCTTCGGGGGCGCGTCGAGTCCGGGGAACCGAGACGACAT							
ACGAGAGCTCAGACACCGATGACGCTTCGGGGGCGCGTCGAGTCCGGGGAACCGAGACGACAT							
ACGAGAGCTCAGACACCGATGACGCTTCGGGGGCGCGTCGAGTCCGGGGAACCGAGACGACAT							
150	160	170	180	190	200		
ACGAGAGCTCAGACACCGATGACGCTTCGGGGGCGCGTCGAGTCCGGGGAACCGAGACGACAT							
ACGAGAGCTCAGACACCGATGACGCTTCGGGGGCGCGTCGAGTCCGGGGAACCGAGACGACAT							
ACGAGAGCTCAGACACCGATGACGCTTCGGGGGCGCGTCGAGTCCGGGGAACCGAGACGACAT							
X							

3. US-09-697-123B-16 (1-208)
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score	=	208	Optimized Score	=	208	Significance	=	1.02
Residue Identity	=	100%	Matches	=	208	Mismatches	=	0
Gaps	=	0	Conservative Substitutions	=			=	0

[illegible][illegible]

ACGAGGCGTACGACCGATGACGTTCCGGGGGGGGTCCAGGTGCGCCGTGGAAACCGACAGACTT
150 160 170 180 190 200 X
ACGAGGCGTACGACCGATGACGTTCCGGGGGGGGTCCAGGTGCGCCGTGGAAACCGACAGACTT
150 160 170 180 190 200 X

4. US-09-697-123B-16 (1-208)
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score	=	207	Optimized Score	=	207	Significance	=	1.00
Residue Identity	=	998	Matches	=	207	Mismatches	=	1
Gaps	=	0	Conservative Substitutions				=	0

[illegible][illegible][illegible]

5. US-09-697-123B-16 (1-208)

US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score	=	186	Optimized Score	=	186	Significance	=	0.60
Residue Identity	=	89%	Matches	=	186	Mismatches	=	22
Gaps	=	0	Conservative Substitutions	=			=	0

X 10 20 30 40 50 60 70
TAAAGAGAGCGCTACGACCTGGCCCCGCTGGTGGCTATAAGTCAACAGAAAGCTGGGCTGCATGTGG
|||||
TCAAGAGAGCGCTACGACCTGGCTGGCTGGGTCAAGTCAACAGAAAGCTGGGCTCAAGCGCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGAGCCATACGTCGTCGACGCTGCACGAAGAAGCGTCGTGGCCACCATCGAATATCTGTCGGCTGC
GCGAGCCATACGTCGTCGACGCTGCACGAAGAAGCGTCGTGGCCACCATCGAATATCTGTCGGCTGC
GCGAGCCATACGTCGTCGACGCTGCACGAAGAAGCGTCGTGGCCACCATCGAATATCTGTCGGCTGC
GCGAGCCATACGTCGTCGACGCTGCACGAAGAAGCGTCGTGGCCACCATCGAATATCTGTCGGCTGC

[illegible]

6. US-09-697-123B-16 (1-208)

US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score	=	185	Optimized Score	=	185	Significance	=	0.58
Residue Identity	=	88%	Matches	=	185	Mismatches	=	23
Gaps	=	0	Conservative Substitutions	=	0			

X 10 20 30 40 50 60 70
 TCACGAGCAACCGTACGACCTGGCCGGTGGGTCCGCTATTAAGTCACACGAAGCTGGGCTGCATGTG
 TCGAGGAGAACCGCTACGACCTGGCCGGTGGGTCCGCTATTAAGTCACACGAAGCTGGGCTGCATGTG
 X 10 20 30 40 50 60 70

GCAGCCATACCAAGCTCGACGCTGACCCGAGGAACACGCTCGCCACCATCGATGATTCGCTTCC
80 90 100 110 120 130 140
GGAGCCATCATCGCGACGCTGACCCGAGGAACACGCTCGCCACCATCGATGATTCGCTTCC
11 21 31 41 51 61 71 81 91 101 111 121 131 141
GCAGCCATACCAAGCTCGACGCTGACCCGAGGAACACGCTCGCCACCATCGATGATTCGCTTCC
80 90 100 110 120 130 140

150	160	170	180	190	200	X
ACGAGGGTCATGAC	AGATGACGTCGGG	GGCGCTCGAG	GTCGCCG	GTGGA	AAACGACGACAT	T
ACGAGGGTCATGAC	AGATGACGTCGGG	GGCGCTCGAG	GTCGCCG	GTGGA	AAACGACGACAT	T
ACGAGGGTCATGAC	AGATGACGTCGGG	GGCGCTCGAG	GTCGCCG	GTGGA	AAACGACGACAT	T
ACGAGGGTCATGAC	AGATGACGTCGGG	GGCGCTCGAG	GTCGCCG	GTGGA	AAACGACGACAT	T
ACGAGGGTCATGAC	AGATGACGTCGGG	GGCGCTCGAG	GTCGCCG	GTGGA	AAACGACGACAT	T
150	160	170	180	190	200	X
ACGAGGGTCATGAC	AGATGACGTCGGG	GGCGCTCGAG	GTCGCCG	GTGGA	AAACGACGACAT	T
ACGAGGGTCATGAC	AGATGACGTCGGG	GGCGCTCGAG	GTCGCCG	GTGGA	AAACGACGACAT	T
ACGAGGGTCATGAC	AGATGACGTCGGG	GGCGCTCGAG	GTCGCCG	GTGGA	AAACGACGACAT	T
ACGAGGGTCATGAC	AGATGACGTCGGG	GGCGCTCGAG	GTCGCCG	GTGGA	AAACGACGACAT	T
ACGAGGGTCATGAC	AGATGACGTCGGG	GGCGCTCGAG	GTCGCCG	GTGGA	AAACGACGACAT	T

7. US-09-697-123B-16 (1-208)

US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score	=	185	Optimized Score	=	185	Significance	=	0.58
Residue Identity	=	88%	Matches	=	185	Mismatches	=	23
Gaps	=	0	Conservative Substitutions	=			=	0

[illegible]

80 90 100 110 120 130 140
 GCGAGCCATACGTCGTGCAGCGTACCGAAGAACGTGTGTGCCACCATCATCTGTGTGGCTTT
 |||||
 GCGATTCGATACACAGCTCCACGCTACCGAAGAACGTGTGTGCCACCATGAGTACCTGGTCCGTCT
 |||||
 80 90 100 110 120 130 140

ACGAGGTCAGACACCATGACCGCTTCGGGGCGCGCTGCAGTGCCTGGTGAACCCGACACAT
150 160 170 180 190 200 X
|||||
ACGAGGTCAGACACCATGACCGCTTCGGGGCGCGCTGCAGTGCCTGGTGAACCCGACACAT
150 160 170 180 190 200 X

8. US-09-697-123B-16 (1-208)

8. US-09-697-123B-16 (1-208)

US-09-697-123b-1 Sequence 15, Application US/09697123b

Initial Score = 184 Optimized Score = 184 Significance = 0.56
Residue Identity = 88% Matches = 184 Mismatches = 24
Gaps = 0 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAAAGCCCTACGACCTGCGCCCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
TCAAGAGAAAGCCCTACGACCTGCGCCCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 10 20 30 40 50 60 70

150 160 170 180 190 200
ACGAGGCTCAGACACGATGACGCTGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
ACGAGGCTCAGACACGATGACGCTGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 150 160 170 180 190 200

9. US-09-697-123b-16 (1-208)
US-09-697-123b-2 Sequence 23, Application US/09697123b

Initial Score = 182 Optimized Score = 182 Significance = 0.52
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAAAGCCCTACGACCTGCGCCCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
TCAAGAGAAAGCCCTACGACCTGCGCCCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 80 90 100 110 120 130 140

150 160 170 180 190 200
ACGAGGCTCAGACACGATGACGCTGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
ACGAGGCTCAGACACGATGACGCTGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 150 160 170 180 190 200

10. US-09-697-123b-16 (1-208)
US-09-697-123b-1 Sequence 19, Application US/09697123b

Initial Score = 182 Optimized Score = 182 Significance = 0.52
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAAAGCCCTACGACCTGCGCCCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
TCAAGAGAAAGCCCTACGACCTGCGCCCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 80 90 100 110 120 130 140

150 160 170 180 190 200
ACGAGGCTCAGACACGATGACGCTGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
ACGAGGCTCAGACACGATGACGCTGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 150 160 170 180 190 200

ACGAGGCTCAGACACGATGACGCTGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
150 160 170 180 190 200 X

11. US-09-697-123b-16 (1-208)
US-09-697-123b-8 Sequence 8, Application US/09697123b

Initial Score = 182 Optimized Score = 182 Significance = 0.52
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAAAGCCCTACGACCTGCGCCCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
TCAAGAGAAAGCCCTACGACCTGCGCCCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 80 90 100 110 120 130 140

150 160 170 180 190 200
ACGAGGCTCAGACACGATGACGCTGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
ACGAGGCTCAGACACGATGACGCTGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 150 160 170 180 190 200

12. US-09-697-123b-16 (1-208)
US-09-697-123b-1 Sequence 1, Application US/09697123b

Initial Score = 182 Optimized Score = 182 Significance = 0.52
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAAAGCCCTACGACCTGCGCCCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
TCAAGAGAAAGCCCTACGACCTGCGCCCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 80 90 100 110 120 130 140

150 160 170 180 190 200
ACGAGGCTCAGACACGATGACGCTGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
ACGAGGCTCAGACACGATGACGCTGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 150 160 170 180 190 200

13. US-09-697-123b-16 (1-208)
US-09-697-123b-2 Sequence 2, Application US/09697123b

Initial Score = 181 Optimized Score = 181 Significance = 0.50
Residue Identity = 87% Matches = 181 Mismatches = 27
Gaps = 0 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAAAGCCCTACGACCTGCGCCCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
TCAAGAGAAAGCCCTACGACCTGCGCCCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
GCGAGCCCATCAGCTGCTGACGCTGACGCGGCGGCGTCCGCTATAGGTCAACAAAGAGCTCGGCGTGCATCTGC
X 80 90 100 110 120 130 140

Initial Score =	144	Optimized Score =	170	Significance =	-0.21
Residue Identity =	80%	Matches	180	Mismatches	28
Gaps =	15	Conservative Substitutions			0

25. US-09-697-123B-16 (1-208)
US-09-607-123B-? Sequence 35 April 2004
US-0000000000

Initial Score	=	19	Optimized Score	=	19	Significance	=	-2.62
Residue Identity	=	100%	Matches	=	19	Mismatches	=	0
Gaps	=	0	Conservative Substitutions	=	0			

X	10	X
TCACGAGAACGCTACGA		
TCACGAGAACGCTACGA		
X	10	X

[illegible]

3. US-09-697-123B-15 (1-208)
US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial score	=	194	Optimized Score	=	194	Significance	=	0.71
Residue Identity	=	93%	Matches	=	194	Mismatches	=	14
Gaps	=	0	Conservative Substitutions	=			=	0

X	10	20	30	40	50	60	70
TCATGAGAAAGCGCTACACACTGGCCCGGTGGGCCGTACAAGGTCAACAAGAAGTGCAGCTGCAGACGGC							
TCAAAGGAAGGCGCTACACACTGGCCCGGTGGGCCGTACAAGGTCAACAAGAAGTGCAGCTGCAGACGGC							
X	10	20	30	40	50	60	70
TCATGAGAAAGCGCTACACACTGGCCCGGTGGGCCGTACAAGGTCAACAAGAAGTGCAGCTGCAGACGGC							
TCAAAGGAAGGCGCTACACACTGGCCCGGTGGGCCGTACAAGGTCAACAAGAAGTGCAGCTGCAGACGGC							

[illegible]

150	160	170	180	190	200	X
ACGAGGCTCAGCCAGCATGACGATCCCGGCGGCGGATGAGTCCGCTGAGACCCACACAT						
150	160	170	180	190	200	X
ACGAGGCTCAGCACACAGATGACCTTCCGCGGCGGACCCAGGATTCCTGGTGAACCGACACAT						
150	160	170	180	190	200	X

4. US-09-697-123B-15 (1-208)
US-09-697-123B-2 Sequence 2, Application US/09697123B

Initial Score	=	194	Optimized Score	=	194	Significance	=	0.71
Residue Identity	=	93%	Matches	=	194	Mismatches	=	14
Gaps	=	0	Conservative Substitutions	=			=	0

TCACGAGACAAGCGCTACGACCCTGGCCGGGTGGCCGTACAGAATCAACAGAAGCTGGGCCACAGCGC
X 10 20 30 40 50 60 70

TCACGAGACAAGCGCTACGACCCTGGCCGGGTGGCCGTACAGAATCAACAGAAGCTGGGCCACAGCGC
10 20 30 40 50 60 70

GTGAGCCGATACCAAGCTTCAGACCTACCGAGGAAGAGCTGCGCCACCAAGATTAAGTTCGGCGGCG
80 90 100 110 120 130 140
TGTGAGGCGATTCAGAGCTTCAGACCTACCGAGGAAGAGCTGCGCCACCAAGATTAAGTTCGGCGGCG
1 10 20 30 40 50 60 70 80 90 100 110 120 130 140
GCAGGCCGATTCAGAGCTTCAGACCTACCGAGGAAGAGCTGCGCCACCAAGATTAAGTTCGGCGGCG
80 90 100 110 120 130 140

150	160	170	180	190	200	
ACGAGCGTACGCGATGACGACCGTCCCGCGCGATGAGTGGCGGTGACAGCAGACAT						X
150	160	170	180	190	200	
ACGAGCGTACGCGCGATGACGCGTCCCGCGCGATGAGTGGCGGTGACAGCAGACAT						X
150	160	170	180	190	200	
ACGAGCGTACGCGCGATGACGCGTCCCGCGCGATGAGTGGCGGTGACAGCAGACAT						X

5. US-09-697-123B-15 (1-208)
US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score	Optimized Score	Significance
191	191	0.65
Residue Identity = 91%	Matches = 191	Mismatches = 17
Gaps = 0	Conservative Substitutions	= 0

x	10	20	30	40	50	60	70
---	----	----	----	----	----	----	----

TC AAGGAGGAGCGCTACGACCTGCGCGCGGTGSGGCGGCTACAAAGTCAACAAGAGCTCGGCTGCACGCGCC
TTC AAGGAGGAGCGCTACGACCTGCGCGCGGTGSGGCGGCTACAAAGTCAACAAGAGCTCGGCTGCACGCGCC
TCAAGGAGGAGCGCTACGACCTGCGCGCGGTGSGGCGGCTACAAAGTCAACAAGAGCTCGGCTGCACGCGCC
X 10 20 30 40 50 60 70

[illegible]

150	160	170	180	190	200	
ACGAGGGTCACCCACGATACCGTCCCCGGGGATCGAGTGGCGGTGAGACCCAGACAT						X
ACGAGGGCCAGACCGCGATACCGCTCCGGGGGTGTGAGAGTGGCGGTGAGACCCAGACAT						
150	160	170	180	190	200	X

6. US-09-697-123B-15 (1-208)
US-09-697-123B-1

Initial score	-	190	Optimized score	-	190	Significance	-	0.63
Residue identity	-	91%	Matches	-	190	Mismatches	-	18
Gaps	-	0	Conservative substitutions	-			-	0

TCAGGAGAAAGGGTACGACCTGGCCCGCGGTGGCCGCTACAAAGTCAACAAAGCTGGCCTGCACGCGG
 X 10 20 30 40 50 60 70
 TCAGGAGAAAGGGTACGACCTGGCCCGCGGTGGCCGCTACAAAGTCAACAAAGCTGGCCTGCACGCGG
 X 10 20 30 40 50 60 70

GTGAGCCGATCAACAGCTGCAGTCCAGAGACAGTGTCGCCACCATTGAGTAATACCTGGTGGCGCTC
80 90 100 110 120 130 140
| | | | | | | | | | | | | | | |
GTTGAGCCGATCAACAGCTGCAGTCCAGAGACAGTGTCGCCACCATTGAGTAATACCTGGTGGCGCTC
90 100 110 120 130 140
| | | | | | | | | | | | | | | |
GCCAAGCCGATCAACAGCTGCAGTCCAGAGACAGTGTCGCCACCATTGAGTAATACCTGGTGGCGCTC
80 90 100 110 120 130 140

	150	160	170	180	190	200	X
ACGAGGCTCAGCCCAACGATGACCCTCCCCGGCGGCATGTGAGTCCGGTGAGACCGAGCAT							
ACGAGGCCAGAACCAAGATGACCTCTCCC GGCGGCCTTCAGTGTCCCGGTGAGGTGAGAGCACT							
150	160	170	180	190	200	X	

7. US-09-697-123B-15 (1-208)

Initial Score	-	188	Optimized Score	-	188	Significance	-	0.59
Residue Identity	-	90%	Matches	-	188	Mismatches	-	20
Gaps	-	0	Conservative Substitutions	-			-	0

X
 10
 20
 30
 40
 50
 60
 70
 TCAAGAGAGAGCGCTACGACCTGGCCCGGGTGGCGCTTACAAAGCTAACAGAGCTTGCGCTGACAGCGG
 X
 10
 20
 30
 40
 50
 60
 70
 TCAGAGAGAGCGCTACGACCTGGCCCGGGTGGCTTACAAAGCTAACAGAGCTTGCGCTGACAGCGG

80 90 100 110 120 130 140
 GTGAGCCGATCACCAGCTCGACGCTACCGAGAGAGAGCTGTGGCCACCATCGAGTACCTGTGGCCCTGC
 |||||
 GTGAGCCGATCACCAGCTCGACGCTACCGAGAGAGAGCTGTGGCCACCATCGAGTACCTGTGGCCCTGC
 |||||
 GTGAGCCGATCACCAGCTCGACGCTACCGAGAGAGAGCTGTGGCCACCATCGAGTACCTGTGGCCCTGC
 80 90 100 110 120 130 140

150	160	170	180	190	200	
ACGAGGGTCAGCCCGATGACCGTCCCGCGGCATGAGTGGCGGTGGAGACCGAGACAT						X
150	160	170	180	190	200	
ATGAGGGTCAGTCGACGATGACCGTCCAGGTGGCGGTGGAGACCGAGTACTGACGACAT						
150	160	170	180	190	200	X

8. US-09-697-123B-15 (1-208)

US-09-697-123B-1 Sequence 10, Application US/09697123B

[illegible]

150 160 170 180 190 200 X
ACGAGGCTACACCCACGATGACCGTCCCGGCGATCGAGTCCCGGTGAGAGCCAGACAT
|||||
ACGAGGCTACACCCACGATGACCGTCCCGGCGATCGAGTCCCGGTGAGAGCCAGACAT
150 160 170 180 190 200 X

14. US-09-697-123B-15 (1-208)
US-09-697-123B-1 Sequence 17, Application US/09697123B

Initial Score = 183 Optimized Score = 183 Significance = 0.49
Residue Identity = 87% Matches = 183 Mismatches = 25
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGCTCAACAGAGCTCGGCTGCACGCG
|||||
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGCTCAACAGAGCTCGGCTGCACGCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCAGTACCTGCTGCGCTGC
|||||
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCAGTACCTGCTGCGCTGC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCTACACCCACGATGACCGTCCCGGCGATCGAGTCCCGGTGAGAGCCAGACAT
|||||
ACGAGGCTACACCCACGATGACCGTCCCGGCGATCGAGTCCCGGTGAGAGCCAGACAT
150 160 170 180 190 200 X

15. US-09-697-123B-15 (1-208)
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.47
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGCTCAACAGAGCTCGGCTGCACGCG
|||||
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGCTCAACAGAGCTCGGCTGCACGCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCAGTACCTGCTGCGCTGC
|||||
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCAGTACCTGCTGCGCTGC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCTACACCCACGATGACCGTCCCGGCGATCGAGTCCCGGTGAGAGCCAGACAT
|||||
ACGAGGCTACACCCACGATGACCGTCCCGGCGATCGAGTCCCGGTGAGAGCCAGACAT
150 160 170 180 190 200 X

16. US-09-697-123B-15 (1-208)
US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score = 177 Optimized Score = 189 Significance = 0.37
Residue Identity = 93% Matches = 195 Mismatches = 10
Gaps = 3 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGCTCAACAGAGCTCGGCTGCACGCG
|||||
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGCTCAACAGAGCTCGGCTGCACGCG
X 10 20 30 40 50 60 70

GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCAGTACCTGCTGCGCTGC
|||||
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCAGTACCTGCTGCGCTGC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCTACACCCACGATGACCGTCCCGGCGATCGAGTCCCGGTGAGAGCCAGACAT
|||||
ACGAGGCTACACCCACGATGACCGTCCCGGCGATCGAGTCCCGGTGAGAGCCAGACAT
150 160 170 180 190 200 X

17. US-09-697-123B-15 (1-208)
US-09-697-123B-1 Sequence 12, Application US/09697123B

Initial Score = 160 Optimized Score = 183 Significance = 0.04
Residue Identity = 90% Matches = 188 Mismatches = 19
Gaps = 1 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGCTCAACAGAGCTCGGCTGCACGCG
|||||
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGCTCAACAGAGCTCGGCTGCACGCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCAGTACCTGCTGCGCTGC
|||||
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCAGTACCTGCTGCGCTGC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCTACACCCACGATGACCGTCCCGGCGATCGAGTCCCGGTGAGAGCCAGACAT
|||||
ACGAGGCTACACCCACGATGACCGTCCCGGCGATCGAGTCCCGGTGAGAGCCAGACAT
150 160 170 180 190 200 X

18. US-09-697-123B-15 (1-208)
US-09-697-123B-4 Sequence 4, Application US/09697123B

Initial Score = 156 Optimized Score = 185 Significance = -0.04
Residue Identity = 91% Matches = 190 Mismatches = 17
Gaps = 1 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGCTCAACAGAGCTCGGCTGCACGCG
|||||
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGCTCAACAGAGCTCGGCTGCACGCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCAGTACCTGCTGCGCTGC
|||||
GTGAGCCGATACACGCTCGACGCTGACCCGAGAGAGCTGCTGCGCACCAGTACCTGCTGCGCTGC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCTACACCCACGATGACCGTCCCGGCGATCGAGTCCCGGTGAGAGCCAGACAT
|||||
ACGAGGCTACACCCACGATGACCGTCCCGGCGATCGAGTCCCGGTGAGAGCCAGACAT
150 160 170 180 190 200 X

19. US-09-697-123B-15 (1-208)
US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 148 Optimized Score = 175 Significance = -0.20
Residue Identity = 84% Matches = 189 Mismatches = 19
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCGGGTGGCGGCTCAAGCTCAACAGAGCTCGGCTGCACGCG
|||||


```
|||||
TCACAGGAGGCGTACGACCTGGCCCGCGGTACAGTCAAGTCAAGTCTGCTGCAACGCCG
X 10 20 30 40 50 60 70
80 90 100 110 120 130 140
GTACAGCGGTATCACAGCTGTGACCTGACCGAGAGAGAGTGTGCTGCGCATGCGTGGCGCTGC
|||||
CGGAGCGCATGACGTCGTACAGCTGAGAGAGAGTGTGCTGCGCATGCGTGGCGCTGC
80 90 100 110 120 130 140
AC-----GAGGTACAGCCAC---GATGACCGTCCCGCGCGCATGAGTGTGAGACCG
|||||
ACCGCCCGGTACGAGTGGCGCGCGCGGTATGATGACGTCCCGCGCGCATGAGTGTGAGACCG
150 160 170 180 190 200 210
X
ACGACAT
|||||
ACGACAT
220 X
```

20. US-09-697-123b-15 (1-208)
US-09-697-123b-1 Sequence 13, Application US/09697123b

Initial Score = 143 Optimized Score = 172 Significance = -0.29
Residue Identity = 81% Matches = 182 Mismatches = 26
Gaps = 15 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCACAGAGAACCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
|||||
TCACAGAGAGAACCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
10 20 30 40 50 60 70
GTACAGCGGTATCACAGCTGTGACCTGACCGAGAGAGTGTGCTGCGCATGCGTGGCGCTGC
|||||
ATCATTCGATATCACAGCTGTGACCTGACCGAGAGAGTGTGCTGCGCATGCGTGGCGCTGC
80 90 100 110 120 130 140
ACGAGC-----GTCAGCCCAAGTACGACCTGGCCCGCGCATGAGTGTGCGAGACCG
|||||
ACCGCCCTCTCAGGCGTGGCGCGCGGTATGACGTCCCGCGCGCATGAGTGTGCGAGAACCG
150 160 170 180 190 200 210
X
ACGACAT
|||||
ACGACAT
220 X
```

21. US-09-697-123b-15 (1-208)
US-09-697-123b-7 Sequence 7, Application US/09697123b

Initial Score = 131 Optimized Score = 170 Significance = -0.53
Residue Identity = 82% Matches = 177 Mismatches = 31
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
|||||
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
10 20 30 40 50 60 70
GTACAGCGGTATCACAGCTGTGACCTGACCGAGAGAGAGTGTGCTGCGCATGCGTGGCGCTGC
|||||
GTACAGCGGTATCACAGCTGTGACCTGACCGAGAGAGAGTGTGCTGCGCATGCGTGGCGCTGC
80 90 100 110 120 130 140
X
GTACAGCGGTATCACAGCTGTGACCTGACCGAGAGAGAGTGTGCTGCGCATGCGTGGCGCTGC
80 90 100 110 120 130 140
```

```
140 150 160 170 180 190 200
GCCTGCACAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGAGTGTGCGCTGACCGACGACAT
|||||
GCCTGCACAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGAGTGTGCGCTGACCGACGACAT
150 160 170 180 190 200 210 X
X 10 20 30 40 50 60 70
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
|||||
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
10 20 30 40 50 60 70
X
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
10 20 30 40 50 60 70
```

22. US-09-697-123b-15 (1-208)
US-09-697-123b-1 Sequence 18, Application US/09697123b

Initial Score = 117 Optimized Score = 162 Significance = -0.80
Residue Identity = 79% Matches = 168 Mismatches = 40
Gaps = 3 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
|||||
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
10 20 30 40 50 60 70
X
GTACAGCGGTATCACAGCTGTGACCTGACCGAGAGAGAGTGTGCTGCGCATGCGTGGCGCTGC
|||||
ACCGCCCGGTACGAGTGGCGCGCGGTATGATGACGTCCCGCGCGCATGAGTGTGAGACCG
80 90 100 110 120 130 140
X
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
|||||
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
150 160 170 180 190 200 210
```

23. US-09-697-123b-15 (1-208)
US-09-697-123b-2 Sequence 21, Application US/09697123b

Initial Score = 112 Optimized Score = 172 Significance = -0.90
Residue Identity = 83% Matches = 179 Mismatches = 29
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
|||||
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
10 20 30 40 50 60 70
X
GTACAGCGGTATCACAGCTGTGACCTGACCGAGAGAGAGTGTGCTGCGCATGCGTGGCGCTGC
80 90 100 110 120 130 140
X
GTACAGCGGTATCACAGCTGTGACCTGACCGAGAGAGAGTGTGCTGCGCATGCGTGGCGCTGC
80 90 100 110 120 130 140
```

24. US-09-697-123b-15 (1-208)
US-09-697-123b-1 Sequence 14, Application US/09697123b

Initial Score = 112 Optimized Score = 173 Significance = -0.90
Residue Identity = 84% Matches = 180 Mismatches = 28
Gaps = 6 Conservative Substitutions = 0

```
X 10 20 30 40 50 60 70
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
|||||
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
10 20 30 40 50 60 70
X
TCACAGAGAGAGCGCTACGACCTGGCCCGCGGTGCGCGCATGACGTCAAGTCAAGAGAGTGTGCGCTGC
10 20 30 40 50 60 70
```


> 0 <
01/10 Inteligenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-697-123b-14.res made by Shanley on Wed 13 Nov 102 14:07:43 PST.

Query sequence being compared: US-09-697-123B-14 (1-214)
Number of sequences searched: 26
Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123B-14 (1-214) with:
File : US09697123B.seq

```

100-
N -
U 50-
M -
B -
E -
R -
O -
F 10-
S -
E 5-
O -
U -
N -
E -
C -
S 0-
SCORE 0 | 24 | 48 | 71 | 95 | 119 | 143 | 166 | 190 | 214 |
STDDEV -2 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |
PARAMETERS
Similarity matrix      Unitary      K-tuple      4
Mismatch penalty      1          Joining penalty 30
Gap penalty           5.00        Window size    207
Gap size penalty      0.33
Cutoff score          12
Randomization group   0
SEARCH STATISTICS
Scores:      Mean      Median      Standard Deviation
            113        113         40.53
Times:      CPU
            00:00:00.00      Total Elapsed
                        00:00:00.00
Number of residues:      5077
Number of sequences searched: 26
Number of scores above cutoff: 25

```

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Length	Score	Int. Opt.	Sig.	Frame
US-09-697-123B-14	Sequence 14, Application	214	214	214	2.49	0

1. US-09-697-123B-1 Sequence 14, Application 214 214 214 2.49 0
The list of other best scores is:

Sequence Name	Description	Length	Score	Int. Opt.	Sig.	Frame
US-09-697-123B-2	Sequence 7, Application	214	172	172	1.46	0
US-09-697-123B-2	Sequence 21, Application	214	171	171	1.43	0
US-09-697-123B-1	Sequence 1, Application	208	136	169	0.57	0
US-09-697-123B-2	Sequence 23, Application	208	132	166	0.47	0
US-09-697-123B-1	Sequence 11, Application	208	129	166	0.39	0
US-09-697-123B-1	Sequence 10, Application	208	126	163	0.32	0
US-09-697-123B-2	Sequence 20, Application	205	122	162	0.22	0
US-09-697-123B-1	Sequence 12, Application	207	120	167	0.17	0
US-09-697-123B-8	Sequence 8, Application	208	118	168	0.12	0
US-09-697-123B-1	Sequence 19, Application	208	116	168	0.07	0
US-09-697-123B-1	Sequence 18, Application	211	115	166	0.05	0
US-09-697-123B-3	Sequence 15, Application	208	112	173	-0.02	0
US-09-697-123B-2	Sequence 3, Application	208	112	168	-0.02	0
US-09-697-123B-2	Sequence 2, Application	208	112	170	-0.02	0
US-09-697-123B-2	Sequence 16, Application	208	110	168	-0.07	0
US-09-697-123B-6	Sequence 6, Application	208	110	168	-0.07	0
US-09-697-123B-5	Sequence 5, Application	208	109	171	-0.10	0
US-09-697-123B-2	Sequence 24, Application	208	109	167	-0.10	0
US-09-697-123B-2	Sequence 22, Application	208	109	169	-0.10	0
US-09-697-123B-1	Sequence 17, Application	208	109	169	-0.10	0
US-09-697-123B-1	Sequence 4, Application	207	97	161	-0.39	0
US-09-697-123B-1	Sequence 13, Application	223	83	152	-0.74	0
US-09-697-123B-9	Sequence 9, Application	223	81	156	-0.79	0
US-09-697-123B-2	Sequence 25, Application	19	19	19	-2.32	0

1. US-09-697-123B-14 (1-214)
US-09-697-123B-1 Sequence 14, Application US/09697123B
Initial Score = 214 Optimized Score = 214 Significance = 2.49
Residue Identity = 100% Matches = 214 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

```

X      10      20      30      40      50      60      70
TCACGAGAGAGCGCTACGACCTGCGCCAGAGGTTGCCGCTTACAGTCAAGTCAAGAGCTGCGCGCGG
|||||
TCACGAGAGAGCGCTACGACCTGCGCCAGAGGTTGCCGCTTACAGTCAAGTCAAGAGCTGCGCGCGG
X      10      20      30      40      50      60      70

```

```

CCGAGTCGCGCGCTACCGCTGACACCGCTGACCGAGGAGGATGTCGCGCCACCATGAGTACCTGTGC
|||||
CCGAGTCGCGCGCTACCGCTGACACCGCTGACCGAGGAGGATGTCGCGCCACCATGAGTACCTGTGC
X      80      90      100      110      120      130      140

```

```

GCGTGCAGAGGCGCCAGGCAACGATGACGGTTCCCGGCGGCGGTGAGGTCGCGAGACCGAGACAT
|||||
GCGTGCAGAGGCGCCAGGCAACGATGACGGTTCCCGGCGGCGGTGAGGTCGCGAGACCGAGACAT
X      150      160      170      180      190      200      210

```

2. US-09-697-123B-14 (1-214)
US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score = 172 Optimized Score = 172 Significance = 1.46
Residue Identity = 80% Matches = 172 Mismatches = 42
Gaps = 0 Conservative Substitutions = 0

```

X      10      20      30      40      50      60      70
TCACGAGAGAGCGCTACGACCTGCGCCAGGAGGTTGCCGCTTACAGTCAAGTCAAGAGCTGCGCGCGG
|||||
TCACGAGAGAGCGCTACGACCTGCGCCAGGAGGTTGCCGCTTACAGTCAAGTCAAGAGCTGCGCGCGG
X      10      20      30      40      50      60      70

```

8. US-09-697-123B-14 (1-214)

[illegible]

80 90 100 110 120 130

150 160 170 180 190 200 210 X
GCCGTGCAGAGGCGCCAGGCAACGATATGAGTTCCTCCGCGCGCTGAGGTGCGCGGTGAGAGACCGAGACAT
|||||
GCCTGCAGAGGCGCTACGACGATGATGACGTCCTCCGCGCGCATCGAGGTGAGAGACCGAGACAT
140 150 160 170 180 190 200 X

14. US-09-697-123B-14 (1-214)
US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 112 Optimized Score = 168 Significance = -0.02
Residue Identity = 81% Matches = 175 Mismatches = 33
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGCGCCAGGCTTGCCGCTTACAAGTCAACAAGAGCTGCGGCTCCGCGG
|||||
TCAAGAGAAAGCGCTACGACCTGCGCCAGGCTTGCCGCTTACAAGTCAACAAGAGCTGCGGCTCCGCGG
10 20 30 40 50 60 70
X

80 90 100 110 120 130 140
CCGAGTGGCGCCGTACCCGCTGCGACCGTACGACGATGTCGTCCGACCATCGAGTACCTGTGC
|||||
GCCATCCG-----ATCACAGCTCCACGCTGACCGAAGAGAGCGTCTGCGCACCATCGAGTACCTGTGC
80 90 100 110 120 130

150 160 170 180 190 200 210 X
GCGTGCAGAGGCGCCAGGCAACGATGATGTCCTCCGCGCGCTCCAGAGTGCCTGCGGTGAGAGACCGAGACAT
|||||
GCTTGCACAGAGGCGCTACGACGATGATGTCCTCCGCGCGCTCCAGAGTGCCTGCGGTGAGAGACCGAGACAT
140 150 160 170 180 190 200 X

15. US-09-697-123B-14 (1-214)
US-09-697-123B-2 Sequence 2, Application US/09697123B

Initial Score = 112 Optimized Score = 170 Significance = -0.02
Residue Identity = 82% Matches = 177 Mismatches = 31
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGCGCCAGGCTTGCCGCTTACAAGTCAACAAGAGCTGCGGCTCCGCGG
|||||
TCAAGAGAAAGCGCTACGACCTGCGCCAGGCTTGCCGCTTACAAGTCAACAAGAGCTGCGGCTCCGCGG
10 20 30 40 50 60 70
X

80 90 100 110 120 130 140
CCGAGTGGCGCCGTACCCGCTGCGACCGTACGACGATGTCGTCCGACCATCGAGTACCTGTGC
|||||
GCATGCCG-----ATCACAGCTCCACGCTGACCGAAGAGAGCGTCTGACCATCGAGTACCTGTGC
80 90 100 110 120 130

150 160 170 180 190 200 210 X
GCCTGCAGAGGCGCCAGGCAACGATGATGTCCTCCGCGCGCTGAGAGTGCCTGCGGTGAGAGACCGAGACAT
|||||
GGTTCACAGAGGCGCTACGACGATGATGTCCTCCGCGCGCTGAGAGTGCCTGCGGTGAGAGACCGAGACAT
140 150 160 170 180 190 200 X

16. US-09-697-123B-14 (1-214)
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 110 Optimized Score = 168 Significance = -0.07
Residue Identity = 81% Matches = 175 Mismatches = 33
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGCGCCAGGCTTGCCGCTTACAAGTCAACAAGAGCTGCGGCTCCGCGG
|||||
TCAAGAGAAAGCGCTACGACCTGCGCCAGGCTTGCCGCTTACAAGTCAACAAGAGCTGCGGCTCCGCGG
10 20 30 40 50 60 70
X

80 90 100 110 120 130 140
CCGAGTGGCGCCGTACCCGCTGCGACCGTACGACGATGTCGTCCGACCATCGAGTACCTGTGC
|||||
GCAGGCC-----CATCACAGTGTGCGACGCTGCGCCAGGCAAGAGAGCTGCGGTGAGAGACCGAGACAT
80 90 100 110 120 130

150 160 170 180 190 200 210 X
GCCGTGCAGAGGCGCCAGGCAACGATATGAGTTCCTCCGCGCGCTGAGGTGCGCGGTGAGAGACCGAGACAT
|||||
GCTTGCACAGAGGCGCTACGACGATGATGTCCTCCGCGCGCTCCAGAGTGCCTGCGGTGAGAGACCGAGACAT
140 150 160 170 180 190 200 X

17. US-09-697-123B-14 (1-214)
US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 110 Optimized Score = 168 Significance = -0.07
Residue Identity = 81% Matches = 175 Mismatches = 33
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGCGCCAGGCTTGCCGCTTACAAGTCAACAAGAGCTGCGGCTCCGCGG
|||||
TCAAGAGAAAGCGCTACGACCTGCGCCAGGCTTGCCGCTTACAAGTCAACAAGAGCTGCGGCTCCGCGG
10 20 30 40 50 60 70
X

80 90 100 110 120 130 140
CCGAGTGGCGCCGTACCCGCTGCGACCGTACGACGATGTCGTCCGACCATCGAGTACCTGTGC
|||||
GCAGGCC-----CATCACAGTGTGCGACGCTGCGCCAGGCAAGAGAGCTGCGGTGAGAGACCGAGACAT
80 90 100 110 120 130

150 160 170 180 190 200 210 X
GCGTGCAGAGGCGCCAGGCAACGATGATGTCCTCCGCGCGCTCCAGAGTGCCTGCGGTGAGAGACCGAGACAT
|||||
GCTTGCACAGAGGCGCTACGACGATGATGTCCTCCGCGCGCTCCAGAGTGCCTGCGGTGAGAGACCGAGACAT
140 150 160 170 180 190 200 X

18. US-09-697-123B-14 (1-214)
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 110 Optimized Score = 168 Significance = -0.07
Residue Identity = 81% Matches = 175 Mismatches = 33
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGCGCCAGGCTTGCCGCTTACAAGTCAACAAGAGCTGCGGCTCCGCGG
|||||
TCAAGAGAAAGCGCTACGACCTGCGCCAGGCTTGCCGCTTACAAGTCAACAAGAGCTGCGGCTCCGCGG
10 20 30 40 50 60 70
X

80 90 100 110 120 130 140
CCGAGTGGCGCCGTACCCGCTGCGACCGTACGACGATGTCGTCCGACCATCGAGTACCTGTGC
|||||
GCAGGCC-----CATCACAGTGTGCGACGCTGCGCCAGGCAAGAGAGCTGCGGTGAGAGACCGAGACAT
80 90 100 110 120 130

150 160 170 180 190 200 210 X
GCCTGCAGAGGCGCCAGGCAACGATGATGTCCTCCGCGCGCTGAGAGTGCCTGCGGTGAGAGACCGAGACAT
|||||
GCTTGCACAGAGGCGCTACGACGATGATGTCCTCCGCGCGCTGAGAGTGCCTGCGGTGAGAGACCGAGACAT
140 150 160 170 180 190 200 X

19. US-09-697-123B-14 (1-214)
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score = 109 Optimized Score = 171 Significance = -0.10
Residue Identity = 83% Matches = 178 Mismatches = 30
Gaps = 6 Conservative Substitutions = 0

20. US-09-697-123B-14 (1-214)
US-09-697-123B-2 sequence 22, Application US/09697123B

Initial Score = 109 Optimized Score = 167 Significance = -0.10
Residue Identity = 81% Matches = 174 Mismatches = 34
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
|||||
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CCGAGTGGCCGCTACCCGCTCGACACGCTGACCGAAGCGGATGGTGGCCACCATCGAGTACCTGGTGC
|||||
AGATGGC-----CCAAACACACGACGACCTGACCGAAGAGACGCTGCTGCCACCATCGAATACCTGGTGC
80 90 100 110 120 130

150 160 170 180 190 200 210 X
GCCTGCGACGAGGCGCCAGGCAACGATGACGGTTCGCGGCGGCGTGCAGAGTGGCGGTGACACCGACACAT
|||||
GCTTGCACGAGGCGGACGCGCATGTAAGTCCCGGTGCGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC
140 150 160 170 180 190 200 X

21. US-09-697-123B-14 (1-214)
US-09-697-123B-1 sequence 17, Application US/09697123B

Initial Score = 109 Optimized Score = 169 Significance = -0.10
Residue Identity = 82% Matches = 176 Mismatches = 32
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
|||||
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CCGAGTGGCCGCTACCCGCTCGACACGCTGACCGAAGCGGATGGTGGCCACCATCGAGTACCTGGTGC
|||||
GCGAGCC-----CATCAGCTGCTCGACGCTGACCGAAGACGCTGCGCACCATCGAATATCTGGTGC
80 90 100 110 120 130

150 160 170 180 190 200 210 X
GCCTGCGACGAGGCGCCAGGCAACGATGACGGTTCGCGGCGGCGTGCAGAGTGGCGGTGACACCGACACAT
|||||
GCTTGCACGAGGCGGACGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT
140 150 160 170 180 190 200 X

22. US-09-697-123B-14 (1-214)
US-09-697-123B-4 sequence 4, Application US/09697123B

Initial Score = 109 Optimized Score = 169 Significance = -0.10
Residue Identity = 82% Matches = 176 Mismatches = 32
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
|||||
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CCGAGTGGCCGCTACCCGCTCGACACGCTGACCGAAGCGGATGGTGGCCACCATCGAGTACCTGGTGC
|||||
-----CGTCCCGATCACGACGCTGACCTGACCGAAGAGACGCTGCTGCCACCATCGAATACCTGGTGC
80 90 100 110 120 130

150 160 170 180 190 200 210 X
GCCTGCGACGAGGCGCCAGGCAACGATGACGGTTCGCGGCGGCGTGCAGAGTGGCGGTGACACCGACACAT
|||||
GCTTGCACGAGGCGGACGCGCATGTAAGTCCCGGTGCGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC
140 150 160 170 180 190 200 X

Initial Score = 97 Optimized Score = 161 Significance = -0.39
Residue Identity = 80% Matches = 173 Mismatches = 34
Gaps = 7 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
|||||
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
X 10 20 30 40 50 60 70

23. US-09-697-123B-14 (1-214)
US-09-697-123B-1 sequence 13, Application US/09697123B

Initial Score = 83 Optimized Score = 152 Significance = -0.74
Residue Identity = 75% Matches = 173 Mismatches = 35
Gaps = 21 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
|||||
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CCGAGTGGCCGCTACCCGCTCGACACGCTGACCGAAGCGGATGGTGGCCACCATCGAGTACCTGGTGC
|||||
GCGATCCG-----ATCAGCAGCTGCGACGCTGACCGAAGAGACGCTGCTGCCACCATCGAATACCTGGTGC
80 90 100 110 120 130

150 160 170 180 190 200 210 X
GCCTGCGACGAGGCGCCAGGCAACGATGACGGTTCGCGGCGGCGTGCAGAGTGGCGGTGACACCGACACAT
|||||
GCTTGCACGAGGCGGACGCGCATGTAAGTCCCGGTGCGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGC
140 150 160 170 180 190 200 X

24. US-09-697-123B-14 (1-214)
US-09-697-123B-9 sequence 9, Application US/09697123B

Initial Score = 81 Optimized Score = 156 Significance = -0.79
Residue Identity = 75% Matches = 172 Mismatches = 36
Gaps = 21 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
|||||
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
X 10 20 30 40 50 60 70

25. US-09-697-123B-14 (1-214)
US-09-697-123B-9 sequence 9, Application US/09697123B

Initial Score = 81 Optimized Score = 156 Significance = -0.79
Residue Identity = 75% Matches = 172 Mismatches = 36
Gaps = 21 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
|||||
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
X 10 20 30 40 50 60 70

26. US-09-697-123B-14 (1-214)
US-09-697-123B-9 sequence 9, Application US/09697123B

Initial Score = 81 Optimized Score = 156 Significance = -0.79
Residue Identity = 75% Matches = 172 Mismatches = 36
Gaps = 21 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
|||||
TCAAGAGAAAGCGCTACGACCTGGCCAGAGGTTGGCCGTTCAAGTCAACAAAGTGGCTGCCGCGG
X 10 20 30 40 50 60 70

```
      80      90     100     110     120     130
150 GCCTGCACGAGGGC-----160
      |||||  |||
140 GGCTGCACGAGCCCGTACGATGAGCCAGCCCGCGTCACTGTCTCCCGGCGCATCGAGGTGCCGCTG
      150     160     170     180     190     200
210 X
AGCCGACGACAT
      |||||
AGCCGACGACAT
      220 X
```

25. US-09-697-123b-14 (1-214)
US-09-697-123b-2 sequence 25, Application US/09697123b

```
Initial Score      = 19 Optimized Score = 19 Significance = -2.32
Residue Identity = 100% Matches = 19 Mismatches = 0
Caps              = 0 Conservative Substitutions = 0

X      10 X
TCAGGAGAGAGCGCTACGA
      |||||
TCAGGAGAGAGCGCTACGA
      10 X
```


FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Query sequence being compared:	US-09-697-123B-13	(1-223)
Number of sequences searched:		26
Number of scores above cutoff:		25

Results of the initial comparison of US-09-697-123B-13 (1-223) with
File : US09697123B.seq

[illegible]

	Uniterary	K-tuple Joining penalty Window size
Similarity matrix	1	4
Mismatch penalty	5.00	30
Gap penalty	0.33	208
Gap size penalty	13	
Cutoff score	0	
Randomization group		

Scores:	Mean	Median	Standard Deviation
	127	138	44.18

Times:	CPU	total Elapsed
--------	-----	---------------

00:00:00.00

Number of residues:

Number of scores above cutoff:

The scores below are sorted by initial score. Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found

Sequence Name	Description	Length	Int. Score	Opt. Score	Sig.	Frame
---------------	-------------	--------	------------	------------	------	-------

Sequence Name	Description	Int. Opt.	Length	Score	Score	Sig.	Frame
---------------	-------------	-----------	--------	-------	-------	------	-------

2.	US-09-697-123B-9	Sequence 9,	Application U	223	195	195	1.54	0
	****	1 standard deviation above mean	****					
3.	US-09-697-123B-8	Sequence 8,	Application U	208	154	181	0.61	0
4.	US-09-697-123B-1	Sequence 17,	Application	208	144	170	0.38	0
5.	US-09-697-123B-1	Sequence 19,	Application	208	143	169	0.36	0
6.	US-09-697-123B-1	Sequence 15,	Application	208	143	172	0.36	0
7.	US-09-697-123B-3	Sequence 3,	Application U	208	142	167	0.34	0
8.	US-09-697-123B-1	Sequence 1,	Application U	208	142	168	0.34	0
9.	US-09-697-123B-2	Sequence 11,	Application	208	141	167	0.32	0
10.	US-09-697-123B-2	Sequence 24,	Application	208	140	168	0.32	0
11.	US-09-697-123B-1	Sequence 10,	Application	208	140	166	0.29	0
12.	US-09-697-123B-2	Sequence 20,	Application U	208	140	166	0.29	0
13.	US-09-697-123B-2	Sequence 20,	Application	205	138	160	0.25	0
14.	US-09-697-123B-1	Sequence 16,	Application	208	137	162	0.23	0
15.	US-09-697-123B-6	Sequence 6,	Application U	208	137	162	0.23	0
16.	US-09-697-123B-2	Sequence 5,	Application U	208	137	162	0.23	0
17.	US-09-697-123B-2	Sequence 23,	Application	208	136	159	0.20	0
18.	US-09-697-123B-2	Sequence 22,	Application	208	136	161	0.20	0
19.	US-09-697-123B-4	Sequence 48,	Application U	207	131	153	0.09	0
20.	US-09-697-123B-1	Sequence 18,	Application	211	116	162	-0.25	0
21.	US-09-697-123B-7	Sequence 12,	Application	207	112	160	-0.34	0
22.	US-09-697-123B-8	Sequence 7,	Application U	214	92	150	-0.79	0
23.	US-09-697-123B-1	Sequence 14,	Application	214	83	152	-1.00	0
	****	1 standard deviation below mean	****					
24.	US-09-697-123B-2	Sequence 21,	Application	214	81	155	-1.04	0
	****	2 standard deviations below mean	****					
25.	US-09-697-123B-2	Sequence 25,	Application	19	19	-2.44	0	0

```

1. US-09-697-123B-13 (1-223)
   US-09-697-123B-1 Sequence 13, Application US/09697123B

Initial Score: - 223 Optimized Score - 223 Significance - 2.17
Residue Identity - 100% Matches - 223 Mismatches - 0
Gaps - 0 Conservative Substitutions - 0

```

X	10	20	30	40	50	60	70
TCAA	GGAAGCGCTACG	AGCTGG	CGCCGCGTACA	AAGTCA	CAAGAAGT	GGCGCTGA	CAACCC
TCAA	GGAAGCGCTACG	AGCTGG	CGCCGCGTACA	AAGTCA	CAAGAAGT	GGCGCTGA	CAACCC
TCAA	GGAAGCGCTACG	AGCTGG	CGCCGCGTACA	AAGTCA	CAAGAAGT	GGCGCTGA	CAACCC
X	10	20	30	40	50	60	70

ATATCGATCAGACACACGAGCGTGCAGAGACGTCGTGCGCCACATCAGTACTGTTGGCTTGG
80 90 100 110 120 130 140
ATATCGATCAGACACACGAGCGTGCAGAGACGTCGTGCGCCACATCAGTACTGTTGGCTTGG
80 90 100 110 120 130 140

150 160 170 180 190 200 210
ACACGGCTCTAGGGTGGACAGGACCCCGTTATGACTCTCCCGGGGGTCAAGTGCCTGTGAAACCG
ACACGGCTCTAGGGTGGACAGGACCCCGTTATGACTCTCCCGGGGGTCAAGTGCCTGTGAAACCG
150 160 170 180 190 200 210
ACACGGCTCTAGGGTGGACAGGACCCCGTTATGACTCTCCCGGGGGTCAAGTGCCTGTGAAACCG

220 X
ACGACAT
|||||
ACGACAT
220 X

2. US-09-697-123B-13 (1-223)
US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 195 Optimized Score = 195 Significance = 1.54

130
 140
 150
 160
 170
 180
 190
 200
 210
 220
 230
 240
 250
 260
 270
 280
 290
 300
 310
 320
 330
 340
 350
 360
 370
 380
 390
 400
 410
 420
 430
 440
 450
 460
 470
 480
 490
 500
 510
 520
 530
 540
 550
 560
 570
 580
 590
 600
 610
 620
 630
 640
 650
 660
 670
 680
 690
 700
 710
 720
 730
 740
 750
 760
 770
 780
 790
 800
 810
 820
 830
 840
 850
 860
 870
 880
 890
 900
 910
 920
 930
 940
 950
 960
 970
 980
 990
 1000

ACGAGG-----GTCAGCCCAACGATGACCGTCCCGCGGCATCAGAGTGCCCGTGGAGACCG
150 160 170 180 190 200

220 X
ACGACAT
1111111
ACGACAT
X

7. US-09-697-123B-13 (1-223)
US-09-697-123B-3 Sequence 3, Application US/09697123B

```
Initial Score = 142 Optimized Score = 167 Significance = 0.34
Residue Identity = 79% Matches = 177 Mismatches = 31
Gaps = 15 Conservative Substitutions = 0
```

X	10	20	30	40	50	60	70
TCAAGGAAACCGCTACGACTGGCCGCCTGCCGTACAAGTCAACAGAAGATGGGCTTCACACCGC 							
TCAAGGAAGCCCTACGACTGGCCGCCTGCCGTACAAGTCAACAGAAGATGGGCTTCACACCGC 							
TCAAGGAAGCCCTACGACTGGCCGCCTGCCGTACAAGTCAACAGAAGATGGGCTTCACACCGC 							
X	10	20	30	40	50	60	70

ATCATCGATCCACCAACAGCAGCGTCCGGAAGAAGACCTCGTCGCGACCATCGAGTACCTGGCTTCC
80 90 100 110 120 130 140
|||||
CGCATCCGATCCACAGTCTACACGCTCGGAGAAACGCTCGTCGACCATCGAGTACCTGGCTTCC
80 90 100 110 120 130 140

[illegible]

220 X
ACGACAT
| | | | |
ACGACAT
X

8. US-09-697-123B-13 (1-223)
US-09-697-123B-1 Sequence 1, Application US/09697123B

Initial Score	=	142	Optimized Score	=	168	Significance	=	0.34
Residue Identity	=	79%	Matches	=	178	Mismatches	=	30
Gaps	=	15	Conservative Substitutions	=			=	0

[illegible]

ATCATCCGATCACCACACGACGCTGACCGAAGAAGACGTCGTCCACCATCGATGTCGTCCTCCGTC
80 90 100 110 120 130 140
|||||
GCGATCCGATCACCACGCTCAGCTGACGGAAGAAAGCTGTCCCATCGATCAGCTGGTGGCCCTGTC
80 90 100 110 120 130 140

[illegible]

220 X
ACGACAT
|||||
ACGACAT
X

9. US-09-697-123B-13 (1-223)
US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score	-	141	Optimized Score	-	167	Significance	-	0.32
Residue Identity	-	79%	Matches	-	177	Mismatches	-	31
Gaps	-	15	Conservative Substitutions	-			-	0

TCAAGGAGAGCGCTACGACCTGAGCGCCGCTGCGCGCTCAAGGTCAACAAAGCGAGCGCTGGAACACG
 10 20 30 40 50 60 70
 TCAAGGAGAGCGCTACGACCTGAGCGCCGCTGCGCGCTCAAGGTCAACAAAGCGAGCGCTGGAACACG
 10 20 30 40 50 60 70

[illegible]

130
 150
 160
 170
 180
 190
 200
 210
 220
 230
 240
 250
 260
 270
 280
 290
 300
 310
 320
 330
 340
 350
 360
 370
 380
 390
 400
 410
 420
 430
 440
 450
 460
 470
 480
 490
 500
 510
 520
 530
 540
 550
 560
 570
 580
 590
 600
 610
 620
 630
 640
 650
 660
 670
 680
 690
 700
 710
 720
 730
 740
 750
 760
 770
 780
 790
 800
 810
 820
 830
 840
 850
 860
 870
 880
 890
 900
 910
 920
 930
 940
 950
 960
 970
 980
 990
 1000
 1010
 1020
 1030
 1040
 1050
 1060
 1070
 1080
 1090
 1100
 1110
 1120
 1130
 1140
 1150
 1160
 1170
 1180
 1190
 1200
 1210
 1220
 1230
 1240
 1250
 1260
 1270
 1280
 1290
 1300
 1310
 1320
 1330
 1340
 1350
 1360
 1370
 1380
 1390
 1400
 1410
 1420
 1430
 1440
 1450
 1460
 1470
 1480
 1490
 1500
 1510
 1520
 1530
 1540
 1550
 1560
 1570
 1580
 1590
 1600
 1610
 1620
 1630
 1640
 1650
 1660
 1670
 1680
 1690
 1700
 1710
 1720
 1730
 1740
 1750
 1760
 1770
 1780
 1790
 1800
 1810
 1820
 1830
 1840
 1850
 1860
 1870
 1880
 1890
 1900
 1910
 1920
 1930
 1940
 1950
 1960
 1970
 1980
 1990
 2000
 2010
 2020
 2030
 2040
 2050
 2060
 2070
 2080
 2090
 2100
 2110
 2120
 2130
 2140
 2150
 2160
 2170
 2180
 2190
 2200
 2210
 2220
 2230
 2240
 2250
 2260
 2270
 2280
 2290
 2300
 2310
 2320
 2330
 2340
 2350
 2360
 2370
 2380
 2390
 2400
 2410
 2420
 2430
 2440
 2450
 2460
 2470
 2480
 2490
 2500
 2510
 2520
 2530
 2540
 2550
 2560
 2570
 2580
 2590
 2600
 2610
 2620
 2630
 2640
 2650
 2660
 2670
 2680
 2690
 2700
 2710
 2720
 2730
 2740
 2750
 2760
 2770
 2780
 2790
 2800
 2810
 2820
 2830
 2840
 2850
 2860
 2870
 2880
 2890
 2900
 2910
 2920
 2930
 2940
 2950
 2960
 2970
 2980
 2990
 3000
 3010
 3020
 3030
 3040
 3050
 3060
 3070
 3080
 3090
 3100
 3110
 3120
 3130
 3140
 3150
 3160
 3170
 3180
 3190
 3200
 3210
 3220
 3230
 3240
 3250
 3260
 3270
 3280
 3290
 3300
 3310
 3320
 3330
 3340
 3350
 3360
 3370
 3380
 3390
 3400
 3410
 3420
 3430
 3440
 3450
 3460
 3470
 3480
 3490
 3500
 3510
 3520
 3530
 3540
 3550
 3560
 3570
 3580
 3590
 3600
 3610
 3620
 3630
 3640
 3650
 3660
 3670
 3680
 3690
 3700
 3710
 3720
 3730
 3740
 3750
 3760
 3770
 3780
 3790
 3800
 3810
 3820
 3830
 3840
 3850
 3860
 3870
 3880
 3890
 3900
 3910
 3920
 3930
 3940
 3950
 3960
 3970
 3980
 3990
 4000
 4010
 4020
 4030
 4040
 4050
 4060
 4070
 4080
 4090
 4100
 4110
 4120
 4130
 4140
 4150
 4160
 4170
 4180
 4190
 4200
 4210
 4220
 4230
 4240
 4250
 4260
 4270
 4280
 4290
 4300
 4310
 4320
 4330
 4340
 4350
 4360
 4370
 4380
 4390
 4400
 4410
 4420
 4430
 4440
 4450
 4460
 4470
 4480
 4490
 4500
 4510
 4520
 4530
 4540
 4550
 4560
 4570
 4580
 4590
 4600
 4610
 4620
 4630
 4640
 4650
 4660
 4670
 4680
 4690
 4700
 4710
 4720
 4730
 4740
 4750
 4760
 4770
 4

220 X
ACGACAT
|||||||
ACGACAT
X

10. US-09-697-123B-13 (1-223)
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score	-	140	Optimized Score	-	168	Significance	-	0.29
Residue Identity	-	79%	Matches	-	178	Mismatches	-	30
Gaps	-	15	Conservative Substitutions	-			-	0

TC AAGAGGAGGCGGTACACAGCTGGCCGCGCTACAGGTTC AACAAGTGGCGCTGACACCGG
 10 20 30 40 50 60 70
 TC AAGAGGAGGCGGTACACAGCTGGCCGCGCTACAGGTTC AACAAGTGGCGCTGACACCGG
 10 20 30 40 50 60 70
 TC AAGAGGAGGCGGTACACAGCTGGCCGCGCTACAGGTTC AACAAGTGGCGCTGACACCGG
 10 20 30 40 50 60 70

ATCATCCGATACCAACCAGCAGCTGCACCGAAGAAGTGTGCCACCATCGATGACTGGTTGGCTGC
| | | | |
AGATTCGCCAACCAACCAACGACCTGACCGAAGAGACGTCTGCCACCATCGAAATACCTGTGGCTGTC

80 90 100 110 120 130 140

150
 ACCACCCCTCTCAGGTCGGCAGCCCGCCCTTATGATCTCCCGCGGGGTGAGGTGGCTGGTAAACCG
 160
 170
 180
 190
 200
 210
 220
 230
 240
 250
 260
 270
 280
 290
 300
 310
 320
 330
 340
 350
 360
 370
 380
 390
 400
 410
 420
 430
 440
 450
 460
 470
 480
 490
 500
 510
 520
 530
 540
 550
 560
 570
 580
 590
 600
 610
 620
 630
 640
 650
 660
 670
 680
 690
 700
 710
 720
 730
 740
 750
 760
 770
 780
 790
 800
 810
 820
 830
 840
 850
 860
 870
 880
 890
 900
 910
 920
 930
 940
 950
 960
 970
 980
 990
 1000
 1010
 1020
 1030
 1040
 1050
 1060
 1070
 1080
 1090
 1100
 1110
 1120
 1130
 1140
 1150
 1160
 1170
 1180
 1190
 1200
 1210
 1220
 1230
 1240
 1250
 1260
 1270
 1280
 1290
 1300
 1310
 1320
 1330
 1340
 1350
 1360
 1370
 1380
 1390
 1400
 1410
 1420
 1430
 1440
 1450
 1460
 1470
 1480
 1490
 1500
 1510
 1520
 1530
 1540
 1550
 1560
 1570
 1580
 1590
 1600
 1610
 1620
 1630
 1640
 1650
 1660
 1670
 1680
 1690
 1700
 1710
 1720
 1730
 1740
 1750
 1760
 1770
 1780
 1790
 1800
 1810
 1820
 1830
 1840
 1850
 1860
 1870
 1880
 1890
 1900
 1910
 1920
 1930
 1940
 1950
 1960
 1970
 1980
 1990
 2000
 2010
 2020
 2030
 2040
 2050
 2060
 2070
 2080
 2090
 2100
 2110
 2120
 2130
 2140
 2150
 2160
 2170
 2180
 2190
 2200
 2210
 2220
 2230
 2240
 2250
 2260
 2270
 2280
 2290
 2300
 2310
 2320
 2330
 2340
 2350
 2360
 2370
 2380
 2390
 2400
 2410
 2420
 2430
 2440
 2450
 2460
 2470
 2480
 2490
 2500
 2510
 2520
 2530
 2540
 2550
 2560
 2570
 2580
 2590
 2600
 2610
 2620
 2630
 2640
 2650
 2660
 2670
 2680
 2690
 2700
 2710
 2720
 2730
 2740
 2750
 2760
 2770
 2780
 2790
 2800
 2810
 2820
 2830
 2840
 2850
 2860
 2870
 2880
 2890
 2900
 2910
 2920
 2930
 2940
 2950
 2960
 2970
 2980
 2990
 3000
 3010
 3020
 3030
 3040
 3050
 3060
 3070
 3080
 3090
 3100
 3110
 3120
 3130
 3140
 3150
 3160
 3170
 3180
 3190
 3200
 3210
 3220
 3230
 3240
 3250
 3260
 3270
 3280
 3290
 3300
 3310
 3320
 3330
 3340
 3350
 3360
 3370
 3380
 3390
 3400
 3410
 3420
 3430
 3440
 3450
 3460
 3470
 3480
 3490
 3500
 3510
 3520
 3530
 3540
 3550
 3560
 3570
 3580
 3590
 3600
 3610
 3620
 3630
 3640
 3650
 3660
 3670
 3680
 3690
 3700
 3710
 3720
 3730
 3740
 3750
 3760
 3770
 3780
 3790
 3800
 3810
 3820
 3830
 3840
 3850
 3860
 3870
 3880
 3890
 3900
 3910
 3920
 3930
 3940
 3950
 3960
 3970
 3980
 3990
 4000
 4010
 4020
 4030
 4040
 4050
 4060
 4070
 4080
 4090
 4100
 4110
 4120
 4130
 4140
 4150
 4160
 4170
 4180
 4190
 4200
 4210
 4220
 4230
 4240
 4250
 4260
 4270
 4280
 4290
 4300
 4310
 4320
 4330
 4340
 4350
 4360
 4370
 4380
 4390
 4400
 4410
 4420
 4430
 4440
 4450
 4460
 4470
 4480
 4490
 4500
 4510
 4520
 4530
 4540
 4550
 4560
 4570
 4580
 4590
 4600
 4610
 4620
 4630
 4640
 4650
 4660
 4670
 4680
 4690
 4700
 4710
 4720
 4730
 4740

220 X
ACGACAT
111111
ACGACAT
X

11. US-09-697/-123B-13 (1-223)
US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score	=	140	Optimized Score	=	166	Significance	=	0.29
Residue Identity	=	78%	Matches	=	176	Mismatches	=	32
Gaps	=	15	Conservative Substitutions	=			=	0

x	10	20	30	40	50	60	70
---	----	----	----	----	----	----	----

TCAAGAGAACCCCTACGACCTGGCCCGCTGCGCCGCTACAGGTCAACAGAAAGCTGGGCTGACACCG
 |||||
 TCAAGAGAACCCCTACGACCTGGCTGCTGGCTGGCTACAGGTCAACAGAAAGCTGGGCTGACACCG
 |||||
 X 10 20 30 40 50 60 70

ATCATTCGGATCACCACACGACCTGACCGAAGAGAGAGCTGCGGACCATGAGTACGCTGCTGCGCTGC
 |||||
 GCCAGCCCATCACCAGCTGACGCTGACCGAAGAGAGAGCTGCGGACCATGAGTACGCTGCTGCGCTGC
 |||||
 X 80 90 100 110 120 130 140

ACCAAGGCTCTTCAGAGGTGGCCAGGCCCCGCTTATGACTGTCCCGGCGGAGTCCGCTGGAACCG
 |||||
 ACAGAGGC-----CAGACCGCATGACGCTCCGCGGCTGCGAGTCCGCTGAGACCG
 |||||
 X 150 160 170 180 190 200 210

220 X
 ACGACAT
 |||||
 ACGACAT
 X

12. US-09-697-123B-13 (1-223)
 US-09-697-123B-2 Sequence 2, Application US/09697123B

Initial Score = 140 Optimized Score = 166 Significance = 0.29
 Residue Identity = 78% Matches = 176 Mismatches = 32
 Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCAAGAGAACCCCTACGACCTGGCCCGCTGCGCCGCTACAGGTCAACAGAAAGCTGGGCTGACACCG
 |||||
 TCAAGAGAACCCCTACGACCTGGCTGCTGGCTGGCTACAGGTCAACAGAAAGCTGGGCTGACACCG
 |||||
 X 10 20 30 40 50 60 70

ATCATTCGGATCACCACACGACCTGACCGAAGAGAGAGCTGCGGACCATGAGTACGCTGCTGCGCTGC
 |||||
 GCAAGCCGATCACCAGCTGACGCTGACCGAAGAGAGAGCTGCGGACCATGAGTACGCTGCTGCGCTGC
 |||||
 X 80 90 100 110 120 130 140

ACCAGCCCTCTCAGAGGTGGCCAGGCCCCGCTTATGACTGTCCCGGCGGAGTCCGCTGGAACCG
 |||||
 ACGAGG-----GTCAGTGGCGGATGACGCTTCCCGGCGGCGGAGTGGCGGTGAGACCG
 |||||
 X 150 160 170 180 190 200 210

220 X
 ACGACAT
 |||||
 ACGACAT
 X

13. US-09-697-123B-13 (1-223)
 US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score = 138 Optimized Score = 160 Significance = 0.25
 Residue Identity = 78% Matches = 175 Mismatches = 30
 Gaps = 18 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCAAGAGAACCCCTACGACCTGGCCCGCTGCGCCGCTACAGGTCAACAGAAAGCTGGGCTGACACCG
 |||||
 TCAAGAGAACCCCTACGACCTGGCTGCTGGCTGGCTACAGGTCAACAGAAAGCTGGGCTGACACCG
 |||||
 X 10 20 30 40 50 60 70

ATCATTCGGATCACCACACGACCTGACCGAAGAGAGAGCTGCGGACCATGAGTACGCTGCTGCGCTGC
 |||||
 GCGAGCCGATCACCAGCTGACGCTGACCGAAGAGAGAGCTGCGGACCATGAGTACGCTGCTGCGCTGC
 |||||
 X 80 90 100 110 120 130 140

80 90 100 110 120 130 140

ACCAAGGCTCTTCAGAGGTGGCCAGGCCCCGCTTATGACTGTCCCGGCGGAGTCCGCTGGAACCG
 |||||
 ACAGAGGC-----CAGCCAGCATGACGCTCCCGGCTGCGAGTCCGCTGAGACCG
 |||||
 X 150 160 170 180 190 200 210

220 X
 ACGACAT
 |||||
 ACGACAT
 X

14. US-09-697-123B-13 (1-223)
 US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 137 Optimized Score = 162 Significance = 0.23
 Residue Identity = 78% Matches = 176 Mismatches = 32
 Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCAAGAGAACCCCTACGACCTGGCCCGCTGCGCCGCTACAGGTCAACAGAAAGCTGGGCTGACACCG
 |||||
 TCAAGAGAACCCCTACGACCTGGCTGCTGGCTGGCTACAGGTCAACAGAAAGCTGGGCTGACACCG
 |||||
 X 10 20 30 40 50 60 70

ATCATTCGGATCACCACACGACCTGACCGAAGAGAGAGCTGCGGACCATGAGTACGCTGCTGCGCTGC
 |||||
 GCGAGCCCATCACCAGCTGACGCTGACCGAAGAGAGAGCTGCGGACCATGAGTACGCTGCTGCGCTGC
 |||||
 X 80 90 100 110 120 130 140

ACCAGCCCTCTCAGAGGTGGCCAGGCCCCGCTTATGACTGTCCCGGCGGAGTCCGCTGGAACCG
 |||||
 AC-----GAGGTACAGACAC--GATGACCGTTCGCGGCGGAGTGGCGGTGAGACCG
 |||||
 X 150 160 170 180 190 200 210

220 X
 ACGACAT
 |||||
 ACGACAT
 X

15. US-09-697-123B-13 (1-223)
 US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 137 Optimized Score = 162 Significance = 0.23
 Residue Identity = 78% Matches = 176 Mismatches = 32
 Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCAAGAGAACCCCTACGACCTGGCCCGCTGCGCCGCTACAGGTCAACAGAAAGCTGGGCTGACACCG
 |||||
 TCAAGAGAACCCCTACGACCTGGCTGCTGGCTGGCTACAGGTCAACAGAAAGCTGGGCTGACACCG
 |||||
 X 10 20 30 40 50 60 70

ATCATTCGGATCACCACACGACCTGACCGAAGAGAGAGCTGCGGACCATGAGTACGCTGCTGCGCTGC
 |||||
 GCGAGCCCATCACCAGCTGACGCTGACCGAAGAGAGAGCTGCGGACCATGAGTACGCTGCTGCGCTGC
 |||||
 X 80 90 100 110 120 130 140

ACCAAGGCTCTTCAGAGGTGGCCAGGCCCCGCTTATGACTGTCCCGGCGGAGTCCGCTGGAACCG
 |||||
 AC-----GAGGTACAGACAC--GATGACCGTTCGCGGCGGAGTGGCGGTGAGACCG
 |||||
 X 150 160 170 180 190 200 210

140 150 160 170 180 190 200 210
GCCCTGACCAACGCGCTTCAGGGTGGCCAGGCCCCCGTTATGACTGTCCCCGGCGGGGTGAGGTGCCCGGTGG

220 x
AAACCGACGACAT
| | | | | | | |
ATGTGACGACAT
210 x

25. US-09-697-123B-13 (1-223)
US-09-697-123B-2 Sequence 25, Application US/09697123B
Initial Score = 19 Optimized Score = 19 Significance = -2.44
Residue Identity = 100% Matches = 19 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0
X 10 X
TCACGAGAGAGCGCTACGA
|||||
TCACGAGAGAGCGCTACGA
X 10 X

Initial Score = 157 Optimized Score = 176 Significance = 0.53
Residue Identity = 87% Matches = 181 Mismatches = 26
Gaps = 1 Conservative Substitutions = 0

```
X
TCAAGAGAAAGCCGTACGACCTGG-TCCGCTCGGCGCTTACAAGTTCACAAAGAAAGCTGCTGACGTCG
|||||
TCAAGAGAAAGCCGTACGACCTGGCGCCCGCTCGCTGCTAATAGTCAACAAAGAAAGCTGCTGACGTCG
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCGAGCCGATCAGCAGTTCGACCGCTGACCGAAGAGATGCTCGCCACCATGCACTGCTGCTGCGCTGC
|||||
GCGAGCCGATCAGCAGTTCGCTGACCGCTGACCGAAGAGATGCTCGCCACCATGCACTGCTGCTGCGCTGC
80 90 100 110 120 130 140
ACGAGGCTCAGACACGATGACCGTTCGCGGCGCGCTGAGGTGCGAAGAACCGACGACAT
150 160 170 180 190 200
ACGAGGCTCAGACACGATGACCGTTCGCGGCGCGCTGAGGTGCGAAGAACCGACGACAT
150 160 170 180 190 200
```

9. US-09-697-123B-12 (1-207)

US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 157 Optimized Score = 176 Significance = 0.53
Residue Identity = 87% Matches = 181 Mismatches = 26
Gaps = 1 Conservative Substitutions = 0

```
X
TCAAGAGAAAGCCGTACGACCTGG-TCCGCTCGGCGCTTACAAGTTCACAAAGAAAGCTGCTGACGTCG
|||||
TCAAGAGAAAGCCGTACGACCTGGCGCCCGCTCGCTGCTAATAGTCAACAAAGAAAGCTGCTGACGTCG
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCGAGCCGATCAGCAGTTCGACCGCTGACCGAAGAGATGCTCGCCACCATGCACTGCTGCTGCGCTGC
|||||
GCGAGCCGATCAGCAGTTCGCTGACCGCTGACCGAAGAGATGCTCGCCACCATGCACTGCTGCTGCGCTGC
80 90 100 110 120 130 140
ACGAGGCTCAGACACGATGACCGTTCGCGGCGCGCTGAGGTGCGAAGAACCGACGACAT
150 160 170 180 190 200
ACGAGGCTCAGACACGATGACCGTTCGCGGCGCGCTGAGGTGCGAAGAACCGACGACAT
150 160 170 180 190 200
```

10. US-09-697-123B-12 (1-207)

US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 157 Optimized Score = 176 Significance = 0.53
Residue Identity = 87% Matches = 181 Mismatches = 26
Gaps = 1 Conservative Substitutions = 0

```
X
TCAAGAGAAAGCCGTACGACCTGG-TCCGCTCGGCGCTTACAAGTTCACAAAGAAAGCTGCTGACGTCG
|||||
TCAAGAGAAAGCCGTACGACCTGGCGCCCGCTCGCTGCTAATAGTCAACAAAGAAAGCTGCTGACGTCG
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCGAGCCGATCAGCAGTTCGACCGCTGACCGAAGAGATGCTCGCCACCATGCACTGCTGCTGCGCTGC
|||||
GCGAGCCGATCAGCAGTTCGCTGACCGCTGACCGAAGAGATGCTCGCCACCATGCACTGCTGCTGCGCTGC
80 90 100 110 120 130 140
ACGAGGCTCAGACACGATGACCGTTCGCGGCGCGCTGAGGTGCGAAGAACCGACGACAT
150 160 170 180 190 200
ACGAGGCTCAGACACGATGACCGTTCGCGGCGCGCTGAGGTGCGAAGAACCGACGACAT
150 160 170 180 190 200
```

11. US-09-697-123B-12 (1-207)

US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 156 Optimized Score = 175 Significance = 0.51
Residue Identity = 86% Matches = 180 Mismatches = 27
Gaps = 1 Conservative Substitutions = 0

```
X
TCAAGAGAAAGCCGTACGACCTGG-TCCGCTCGGCGCTTACAAGTTCACAAAGAAAGCTGCTGACGTCG
|||||
TCAAGAGAAAGCCGTACGACCTGGCGCCCGCTCGCTGCTAATAGTCAACAAAGAAAGCTGCTGACGTCG
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCGAGCCGATCAGCAGTTCGACCGCTGACCGAAGAGATGCTCGCCACCATGCACTGCTGCTGCGCTGC
|||||
GCGAGCCGATCAGCAGTTCGCTGACCGCTGACCGAAGAGATGCTCGCCACCATGCACTGCTGCTGCGCTGC
80 90 100 110 120 130 140
ACGAGGCTCAGACACGATGACCGTTCGCGGCGCGCTGAGGTGCGAAGAACCGACGACAT
150 160 170 180 190 200
ACGAGGCTCAGACACGATGACCGTTCGCGGCGCGCTGAGGTGCGAAGAACCGACGACAT
150 160 170 180 190 200
```

12. US-09-697-123B-12 (1-207)

US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 155 Optimized Score = 178 Significance = 0.48
Residue Identity = 87% Matches = 183 Mismatches = 24
Gaps = 1 Conservative Substitutions = 0

```
X
TCAAGAGAAAGCCGTACGACCTGG-TCCGCTCGGCGCTTACAAGTTCACAAAGAAAGCTGCTGACGTCG
|||||
TCAAGAGAAAGCCGTACGACCTGGCGCCCGCTCGCTGCTAATAGTCAACAAAGAAAGCTGCTGACGTCG
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCGAGCCGATCAGCAGTTCGACCGCTGACCGAAGAGATGCTCGCCACCATGCACTGCTGCTGCGCTGC
|||||
GCGAGCCGATCAGCAGTTCGCTGACCGCTGACCGAAGAGATGCTCGCCACCATGCACTGCTGCTGCGCTGC
80 90 100 110 120 130 140
ACGAGGCTCAGACACGATGACCGTTCGCGGCGCGCTGAGGTGCGAAGAACCGACGACAT
150 160 170 180 190 200
ACGAGGCTCAGACACGATGACCGTTCGCGGCGCGCTGAGGTGCGAAGAACCGACGACAT
150 160 170 180 190 200
```

13. US-09-697-123B-12 (1-207)

US-09-697-123B-8 Sequence 8, Application US/09697123B

Initial Score = 153 Optimized Score = 174 Significance = 0.44
Residue Identity = 86% Matches = 179 Mismatches = 28
Gaps = 1 Conservative Substitutions = 0

```
X
TCAAGAGAAAGCCGTACGACCTGG-TCCGCTCGGCGCTTACAAGTTCACAAAGAAAGCTGCTGACGTCG
|||||
TCAAGAGAAAGCCGTACGACCTGGCGCCCGCTCGCTGCTAATAGTCAACAAAGAAAGCTGCTGACGTCG
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCGAGCCGATCAGCAGTTCGACCGCTGACCGAAGAGATGCTCGCCACCATGCACTGCTGCTGCGCTGC
|||||
GCGAGCCGATCAGCAGTTCGCTGACCGCTGACCGAAGAGATGCTCGCCACCATGCACTGCTGCTGCGCTGC
80 90 100 110 120 130 140
ACGAGGCTCAGACACGATGACCGTTCGCGGCGCGCTGAGGTGCGAAGAACCGACGACAT
150 160 170 180 190 200
ACGAGGCTCAGACACGATGACCGTTCGCGGCGCGCTGAGGTGCGAAGAACCGACGACAT
150 160 170 180 190 200
```

	150	160	170	180	190	200	
ACGAGGGCCACACCGATGACCGTTCCCGGGGACCGAGCTGCCGTGGAGACGACGACAT							X
ACGAGGGCCACACCGATGACCGTTCCCGGGGACCGAGCTGCCGTGGAGACGACGACAT							
ACGAGGGCCACACCGATGACCGTTCCCGGGGACCGAGCTGCCGTGGAGACGACGACAT							
ACGAGGGCCACACCGATGACCGTTCCCGGGGACCGAGCTGCCGTGGAGACGACGACAT							
ACGAGGGCCACACCGATGACCGTTCCCGGGGACCGAGCTGCCGTGGAGACGACGACAT							
150	160	170	180	190	200		X

14. US-09-697-123B-12 (1-207)
US-09-697-123B-2 Sequence 23, Application US/09697123B

Initial Score	-	152	Optimized Score	-	177	Significance	-	0.41
Residue Identity	-	87%	Matches	-	182	Mismatches	-	25
Gaps	-	1	Conservative Substitutions	-			-	0

X 10 20 30 40 50 60 70
TC AAGCAAGCGCTACGACCTGG-TCGCGTGGCGGTTAC AAGTC AACAAGCTCGGTG AAGCTGC
TC AAGGAGAAAGCGCTACGACCTGGCGGGGTTGGTTAC AAGTC AACAAGCTCGGTG AAGCTGC
TC AAGGAGAAAGCGCTACGACCTGGCGGGGTTGGTTAC AAGTC AACAAGCTCGGTG AAGCTGC

CGGACCCGCACCACTTTCAGACCTTGACCAGAAAGATGTGTCGCCACCATGATACCTGGTTCCGCT
 |||||
 GTGACCGCATTACACTTGACGCTGACCGAAGAGAGACTGTGCCACATGAGTAGCTGGTCCGCT

150	160	170	180	190	200	X
ACAGGGCCAGAC	CACAGGAGAC	CTTCCGGGG	CACCGAGTG	CCGGTGAGAC	CCAGCAGCAT	
150	160	170	180	190	200	X
ATAGGCTCAGT	GCAGGAGAC	CTTCCAGGT	GGGTCGAGT	GCACATGAT	CTCAGCAGCAT	

15. US-09-697-123B-12 (1-207)
US-09-697-123B-1 Sequence 17, Application US/09697123B

Initial Score	=	150	Optimized Score	=	172	Significance	=	0.37
Residue Identity	=	85%	Matches	=	177	Mismatches	=	30
Gaps	=	1	Conservative Substitutions	=			=	0

[illegible]

GGGAGCCGATACCAAGTTTCAGCCGTACCGGAAGAGATGTGTCACCAATCGAATACCTGTTGGCTGC
 |||||
 CGTCCCGATACACACACACACTTCACCGAAGAGAGACTGTGTCACCAATCGAATACCTGTTGGCTGC
 80 90 100 110 120 130 140

ACGAGGGCCAGACACGATGACCGTCCCGGGCCACCGAGTCCCGTGAGACCGAGACAT
 150 160 170 180 190 200 X
 ACGAGGGCCACACGACGATGACCGTCCCGGGCCAGTCCCGTGAGAACCGAGACAT
 150 160 170 180 190 200 X

16. US-09-697-123B-12 (1-207)
US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score	=	144	Optimized Score	=	172	Significance	=	0.23
Residue Identity	=	87%	Matches	=	183	Mismatches	=	21
Caps	=	4	Conservative Substitutions				=	0

10 20 30 40 50 60 70
 TCAAGGAAAGCCGTACGACGTGG-TGCGTGGCGCCGTTACAAAGTCACAAAGAGCTGCTGAACCTG
 80 90 100 110 120 130 140
 TCAAGGAAAGCCGTACGACCTTGGCGCGTGTGGCCCTACAAAGCTACAGAGAGAGCTGGCTGCACGGCG

GCAGGCCGATCACACAGTTTCGACGCTGACCCAAAGATGTCGTCCCAACCATCGAGTACCTGGTTTCGGGTGC
GCGGCCGATCACACAGTTTCGACGCTGACCCAAAGATGTCGTCCCAACCATCGAGTACCTGGTTTCGGGTGC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
 ACGAGGGCCAGACCATGATGACGTCCTCCGGGGGAGCAGGATGCCGTGAGACCCACACAT
 150 160 170 180 190 200 X
 ACGAGGGCCAGACCATGACGTCCTCCGGGATGAGGTGCGGTGAGACCCACACAT
 150 160 170 180 190 200 X

17. US-09-697-123B-12 (1-207)
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score	-	144	Optimized Score	-	167	Significance	-	0.2
Residue Identity	-	82%	Matches	-	172	Mismatches	-	3
Gaps	-	1	Conservative Substitutions	-			-	

[illegible]

GGCAGCCGATCACCAGTTTCGACGCTGACCCCAAGAGATGTCTGCGCCACCATGAGTACCTGGTTCGGCGTGC
 80 90 100 110 120 130 140
 AGAATTCGCGGCACACACACGACCTGACCGCAGAGAGACGTGTGGCCACCATGGAATTACTGTGTCCGCTGC
 80 90 100 110 120 130 140

150	160	170	180	190	200	X
ACGAGGGCCAGACACGATGACCGTCCGGCGGCGACGAGGTGCGGTGAGACCGACACAT						
ACGAGGGGACCGCCACGATGAAGTCCCGGTGGCGTTCGAGTCCCGGTGAGACCGACGACAT						
150	160	170	180	190	200	X

18. US-09-697-123B-12 (1-207)
US-09-697-123B-4 Sequence 4, Application US/09697123B

Initial Score	-	122	Optimized Score	-	176	Significance	-	-0.28
Residue Identity	-	89%	Matches	-	186	Mismatches	-	20
Gaps	-	2	Conservative Substitutions	-			-	0

[illegible]

GGGAGCCGATATACCACTTTGCAGCGCTTACCGAAGAGATGTCGTCCACCATTGATCTGTTCCGGCTGC
80 90 100 110 120 130 140
GGGATCCGATATACCACTTGCAGCGCTTACCGAAGAGATGTCGTCCACCATTGATCTGTTCCGGCTGC
90 100 110 120 130 140

	150	160	170	180	190	200	
ACGAGGGCCAGACACACGATGACGTTCCCGGCCGACCGAGGTCGCGGTGAGACCGACGACAT							X
ACGAGGGTCAGACACACGATGA-CGTTCCGGCGGGGACCGAGTTCCGCTGAGACACCGACGACAT							
150	160	170	180	190	200		X

19. US-09-697-123B-12 (1-207)
US-09-697-123B-1 Sequence 14, Application US/09697123B

Initial Score	=	120	Optimized Score	=	167	Significance	=	-0.32
Residue Identity	=	83%	Matches	=	179	Mismatches	=	28
Gaps	=	7	Conservative Substitutions	=			=	0

X
10 20 30 40 50 60 70
TCAAGAGAACCGCTACGACCTGTC-GCGTCGGCGTTACAAAGTCACAAAAAGCTCGGCTGTGAACGTCG

```
|||||
TCAAGAGAACCGCTACGACCTGGCCAGGGTGGCCGTTGACAGGTCAACAAAGAACTGGGGCTGCCGCGG
X      10      20      30      40      50      60      70
GCGAGCCG-----ATCACCAGTTGACCGCTGACCGAAGAGATGTCTGCGCCACCATGAGTACCTGGTTC
|||||
CCGAGTGGCGGCTGACCGCTCGACCGCTGACCGAAGGGGATGTCTGCGCCACCATGAGTACCTGGTTC
80      90      100     110     120     130
140     150     160     170     180     190     200
X
GCGTGCACAGAGGGCGACACCGATGACCGCTCCGCGGACACCGAGTGGCGGAGAGACGACAT
|||||
GCTGCAAGAGGCGCCAGCAACGATGACGCTCCGCGGCGGTGAGAGCGCGGAGACCGACGACAT
150     160     170     180     190     200
X
```

20. US-09-697-123b-12 (1-207)
US-09-697-123b-1 Sequence 13, Application US/09697123B

Initial Score = 112 Optimized Score = 160 Significance = -0.51
Residue Identity = 80% Matches = 179 Mismatches = 28
Gaps = 16 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAGAGCGCTACGACCTGG-TGCGTGGCGCGCTTACAAAGTCAACAAAAGCTGGTCTGAAGTGG
|||||
TCAAGAGAGAGCGCTACGACCTGGCGCGCGCTGCGCTACAAAGTCAACAAAAGTGGGCTGAACACCG
X      10      20      30      40      50      60      70
GCGAGCGGATCCACAGTTCGACGCTGACCGTGAAGAGAGATGTCTGCGCCACCATGAGTACCTGGTTC
|||||
ATCATCGATCCACCGACGACGCTGACCGAAGAGAGCTGCTCCGACCATGAGTACCTGGTTCGCTTC
80      90      100     110     120     130     140
AC-----GAGGGCGACAGACCAC--GATGACCGTTCGCCGCGGACACGAGTGGCGGAGAGACCG
|||||
ACACCGCTCTCTAGGGTGGCGAAGCGCCGTTATGACTGTCTCCCGGGGGGTGAGGTGCCGTTGGAACCG
150     160     170     180     190     200     210
```

X
ACGACAT
|||||
ACGACAT
220 X

21. US-09-697-123b-12 (1-207)
US-09-697-123b-9 Sequence 9, Application US/09697123B

Initial Score = 112 Optimized Score = 164 Significance = -0.51
Residue Identity = 80% Matches = 179 Mismatches = 28
Gaps = 16 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAGAGCGCTACGACCTGG-TGCGTGGCGCGCTTACAAAGTCAACAAAAGCTGGTCTGAAGTGG
|||||
TCAAGAGAGAGCGCTACGACCTGGCGCGCGCTGCGCTACAAAGTCAACAAAAGCTGGTCTGACGCGCG
X      10      20      30      40      50      60      70
GCGAGCGGATCCACAGTTCGACGCTGACCGTGAAGAGAGATGTCTGCGCCACCATGAGTACCTGGTTC
|||||
GCGAGCGGATCCACAGTTCGACGCTGACCGTGAAGAGAGATGTCTGCGCCACCATGAGTACCTGGTTC
80      90      100     110     120     130     140
```

```
150     160     170     180     190     200
ACGAGGCGCAGA-----CCAGATGACCGTTCGCCGCGGACCGAGGTGGCGGTGAGAGACCG
|||||
ACGAGGCGGCTACGAGTGGCGGCGCGCTGATGATGTCTCCGCGGCGCATGAGGTGGCGGTGAGAGACCG
150     160     170     180     190     200     210
```

X
ACGACAT
|||||
ACGACAT
220 X

22. US-09-697-123b-12 (1-207)
US-09-697-123b-2 Sequence 21, Application US/09697123B

Initial Score = 106 Optimized Score = 158 Significance = -0.64
Residue Identity = 79% Matches = 170 Mismatches = 37
Gaps = 7 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAGAGCGCTACGACCTGG-TGCGTGGCGCGCTTACAAAGTCAACAAAAGCTGGTCTGAAGTGG
|||||
TCAAGAGAGAGCGCTACGACCTGGCGCGCGCTGCGCTACAAAGTCAACAAAAGCTGGGCTGGCGCA
X      10      20      30      40      50      60      70
```

```
GCGAGCC-----GATCACCAAGTTGACGCGTGAACCGAAGAGATGTCTGCGCCACCATGAGTACCTGGTTC
|||||
CCATCGGCTCAGGTGACCAACACCTTCAACCGAGAGAGCTGCTCCGACCATGAGTACCTGGTTC
80      90      100     110     120     130     140
140     150     160     170     180     190     200
X
GCGTGCACAGAGGGCGACACCGATGACCGCTCCGCGGCGGCTGCGCGGAGTGGCGGAGACCGACAT
|||||
GCTGCAAGAGGCGCCAGACGACGATGACCGCCCGCGGCGGTGAGAGTGTGAGACAT
150     160     170     180     190     200     210
X
```

23. US-09-697-123b-12 (1-207)
US-09-697-123b-7 Sequence 7, Application US/09697123B

Initial Score = 101 Optimized Score = 158 Significance = -0.76
Residue Identity = 79% Matches = 170 Mismatches = 37
Gaps = 7 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAGAGCGCTACGACCTGG-TGCGTGGCGCGCTTACAAAGTCAACAAAAGCTGGTCTGAAGTGG
|||||
TCAAGAGAGAGCGCTACGACCTGGCGCGCGCTGCGCTACAAAGTCAACAAAAGCTGGTCTGGCGGTG
X      10      20      30      40      50      60      70
```

```
GCGAGCC-----GATCACCAAGTTGACGCGTGAACCGAAGAGATGTCTGCGCCACCATGAGTACCTGGTTC
|||||
CCAAACCGGCTCTGGTGAATGCGCACCGCTTACCGAAGAGAGCTGCTCCGACCATGAGTACCTGGTTC
80      90      100     110     120     130     140
140     150     160     170     180     190     200
X
GCGTGCACAGAGGGCGACACCGATGACCGCTCCGCGGCGGCTGCGCGGAGTGGCGGAGACCGACAT
|||||
GCGTGCACAGAGGGCGACACCGATGACCGCGCGGCGGCTGAGGTGCGGTGAGGTGCGACGACAT
150     160     170     180     190     200     210
X
```

24. US-09-697-123b-12 (1-207)
US-09-697-123b-1 Sequence 18, Application US/09697123B

Initial Score = 98 Optimized Score = 155 Significance = -0.83
Residue Identity = 78% Matches = 166 Mismatches = 41
Gaps = 4 Conservative Substitutions = 0

```
X      10      20      30      40      50      60      70
TCAAGAGAGAGCGCTACGACCTGG-TGCGTGGCGCGCTTACAAAGTCAACAAAAGCTGGTCTGAAGTGG
|||||
TCAAGAGAGAGCGCTACGACCTGGCGCGCGCTGAGTACAAAGTCAACAAAAGCTGGGCTATCACCGAGA
X      10      20      30      40      50      60      70
```


> 0 <
01 10 IntellGenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-697-123b-11.res made by shanley on Wed 13 Nov 102 14:06:36-PST.

Query sequence being compared: US-09-697-123B-11 (1-208)
Number of sequences searched: 26
Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123B-11 (1-208) with:
File : US09697123B.seq

```

100-
N -
U -
M -
B -
E -
R -
O -
F -
S -
E -
U -
N -
C -
S -
0 -
SCORE 0 23 46 69 92 116 139 162 185 208
STDEV -2 -1 0 1 1 1 1 1 1 1

```

PARAMETERS

Similarity matrix Unitary K-tuple
Mismatch penalty 1 Joining penalty 4
Gap penalty 5.00 Window size 30
Gap size penalty 0.33
Cutoff score 12
Randomization group 0

SEARCH STATISTICS

Scores: Mean Median Standard Deviation
157 181 50.23
Times: CPU Total Elapsed
00:00:00.00 00:00:00.00

Number of residues: 5077
Number of sequences searched: 26
Number of scores above cutoff: 25

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.	Length	Score	Sig.	Frame

1. US-09-697-123B-1 Sequence 11, Application 208 208 1.02 0
The list of other best scores is:

Sequence Name	Description	Length	Score	Init. Opt.	Sig.	Frame
2. US-09-697-123B-1 Sequence 10, Application		208	205	205	0.96	0
3. US-09-697-123B-1 Sequence 15, Application		208	191	191	0.68	0
4. US-09-697-123B-2 Sequence 2, Application U		208	189	189	0.64	0
5. US-09-697-123B-1 Sequence 1, Application U		208	189	189	0.64	0
6. US-09-697-123B-1 Sequence 19, Application		208	187	187	0.60	0
7. US-09-697-123B-1 Sequence 16, Application		208	185	185	0.56	0
8. US-09-697-123B-6 Sequence 6, Application U		208	185	185	0.56	0
9. US-09-697-123B-5 Sequence 5, Application U		208	185	185	0.56	0
10. US-09-697-123B-3 Sequence 3, Application U		208	185	185	0.56	0
11. US-09-697-123B-2 Sequence 22, Application		208	184	184	0.54	0
12. US-09-697-123B-8 Sequence 8, Application U		208	184	184	0.54	0
13. US-09-697-123B-1 Sequence 17, Application		208	181	181	0.48	0
14. US-09-697-123B-2 Sequence 23, Application		208	180	180	0.46	0
15. US-09-697-123B-2 Sequence 24, Application		208	177	177	0.40	0
16. US-09-697-123B-2 Sequence 20, Application		205	166	177	0.18	0
17. US-09-697-123B-1 Sequence 12, Application		207	155	178	-0.04	0
18. US-09-697-123B-4 Sequence 4, Application U		207	151	177	-0.12	0
19. US-09-697-123B-1 Sequence 13, Application		223	141	167	-0.32	0
20. US-09-697-123B-9 Sequence 9, Application U		223	140	168	-0.34	0
21. US-09-697-123B-2 Sequence 21, Application		214	138	159	-0.38	0
22. US-09-697-123B-1 Sequence 14, Application		214	129	166	-0.56	0
23. US-09-697-123B-1 Sequence 18, Application		211	121	166	-0.72	0
24. US-09-697-123B-7 Sequence 7, Application U		214	108	168	-0.98	0
25. US-09-697-123B-2 Sequence 25, Application		19	19	19	-2.75	0

1. US-09-697-123B-11 (1-208)

US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 208 Optimized Score = 208 Significance = 1.02
Residue Identity = 100% Matches = 208 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X	10	20	30	40	50	60	70
TCAAGGAGAGCGCCTACGACCTGCGCCGGGTGGCGCGGTACAGTTCACAGAACGCTGCGCTGAACCGC							
TCAAGGAGAGCGCCTACGACCTGCGCCGGGTGGCGCGGTACAGTTCACAGAACGCTGCGCTGAACCGC							
X	10	20	30	40	50	60	70
GCCAGGCCCATCACCGCTGACGCTGACCGAGAGAACGTCGTGCGCACCATGCAATACCTGCTCGCTTGC							
GCCAGGCCCATCACCGCTGACGCTGACCGAGAGAACGTCGTGCGCACCATGCAATACCTGCTCGCTTGC							

150	160	170	180	190	200	X
ACGAGGGCCACGACCGCGATGACCGCGCGGTGTCGAGGTGTCGAGACGACAT						
ACGAGGGCCACGACCGCGATGACCGCGGTGTCGAGGTGTCGAGACGACAT						
150	160	170	180	190	200	X

2. US-09-697-123B-11 (1-208)

US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 205 Optimized Score = 205 Significance = 0.96
Residue Identity = 98% Matches = 205 Mismatches = 3
Gaps = 0 Conservative Substitutions = 0

X	10	20	30	40	50	60	70
TCAAGGAGAGCGCCTACGACCTGCGCCGGGTGGCGCGGTACAGTTCACAGAACGCTGCGCTGAACCGC							
TCAAGGAGAGCGCCTACGACCTGCGCCGGGTGGCGCGGTACAGTTCACAGAACGCTGCGCTGAACCGC							
X	10	20	30	40	50	60	70

Initial Score = 185 Optimized Score = 185 Significance = 0.56
Residue Identity = 88% Matches = 185 Mismatches = 23
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACAGACCTGCGCCCGGCGGCGGCTCAAGGTCAACAAGAGCTGCGCTGAACCGCG
TCAAGAGAAAGCGCTACAGACCTGCGCCCGGCGGCGGCTCAAGGTCAACAAGAGCTGCGCTGAACCGCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCCAGCCCATACACCTGACGCTGACCGAAGAGCTGCTGCGCACATCGAATGCTGCTGCGCTGCG
GCCAGCCCATACACCTGACGCTGACCGAAGAGCTGCTGCGCACATCGAATGCTGCTGCGCTGCG
X 10 20 30 40 50 60 70

150 160 170 180 190 200
ACGAGGGCCAGACCGGCGATGACCGCTCGCGGCGGCGGCTGAGAGTCCGCTGAGACCGACAT X
ACGAGGGCCAGACCGGCGATGACCGCTCGCGGCGGCGGCTGAGAGTCCGCTGAGACCGACAT X
X 150 160 170 180 190 200

9. US-09-697-123B-11 (1-208)
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.56
Residue Identity = 88% Matches = 185 Mismatches = 23
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACAGACCTGCGCCCGGCGGCGGCTCAAGGTCAACAAGAGCTGCGCTGAACCGCG
TCAAGAGAAAGCGCTACAGACCTGCGCCCGGCGGCGGCTCAAGGTCAACAAGAGCTGCGCTGAACCGCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCCAGCCCATACACCTGACGCTGACCGAAGAGCTGCTGCGCACATCGAATGCTGCTGCGCTGCG
GCCAGCCCATACACCTGACGCTGACCGAAGAGCTGCTGCGCACATCGAATGCTGCTGCGCTGCG
X 80 90 100 110 120 130 140

150 160 170 180 190 200
ACGAGGGCCAGACCGGCGATGACCGCTCGCGGCGGCGGCTGAGAGTCCGCTGAGACCGACAT X
ACGAGGGCCAGACCGGCGATGACCGCTCGCGGCGGCGGCTGAGAGTCCGCTGAGACCGACAT X
X 150 160 170 180 190 200

10. US-09-697-123B-11 (1-208)
US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.56
Residue Identity = 88% Matches = 185 Mismatches = 23
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACAGACCTGCGCCCGGCGGCGGCTCAAGGTCAACAAGAGCTGCGCTGAACCGCG
TCAAGAGAAAGCGCTACAGACCTGCGCCCGGCGGCGGCTCAAGGTCAACAAGAGCTGCGCTGAACCGCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCCAGCCCATACACCTGACGCTGACCGAAGAGCTGCTGCGCACATCGAATGCTGCTGCGCTGCG
GCCAGCCCATACACCTGACGCTGACCGAAGAGCTGCTGCGCACATCGAATGCTGCTGCGCTGCG
X 80 90 100 110 120 130 140

150 160 170 180 190 200
ACGAGGGCCAGACCGGCGATGACCGCTCGCGGCGGCGGCTGAGAGTCCGCTGAGACCGACAT X
ACGAGGGCCAGACCGGCGATGACCGCTCGCGGCGGCGGCTGAGAGTCCGCTGAGACCGACAT X
X 150 160 170 180 190 200

11. US-09-697-123B-11 (1-208)
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 184 Optimized Score = 184 Significance = 0.54
Residue Identity = 88% Matches = 184 Mismatches = 24
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACAGACCTGCGCCCGGCGGCGGCTCAAGGTCAACAAGAGCTGCGCTGAACCGCG
TCAAGAGAAAGCGCTACAGACCTGCGCCCGGCGGCGGCTCAAGGTCAACAAGAGCTGCGCTGAACCGCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCCAGCCCATACACCTGACGCTGACCGAAGAGCTGCTGCGCACATCGAATGCTGCTGCGCTGCG
GCCAGCCCATACACCTGACGCTGACCGAAGAGCTGCTGCGCACATCGAATGCTGCTGCGCTGCG
X 80 90 100 110 120 130 140

150 160 170 180 190 200
ACGAGGGCCAGACCGGCGATGACCGCTCGCGGCGGCGGCTGAGAGTCCGCTGAGACCGACAT X
ACGAGGGCCAGACCGGCGATGACCGCTCGCGGCGGCGGCTGAGAGTCCGCTGAGACCGACAT X
X 150 160 170 180 190 200

12. US-09-697-123B-11 (1-208)
US-09-697-123B-8 Sequence 8, Application US/09697123B

Initial Score = 184 Optimized Score = 184 Significance = 0.54
Residue Identity = 88% Matches = 184 Mismatches = 24
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACAGACCTGCGCCCGGCGGCGGCTCAAGGTCAACAAGAGCTGCGCTGAACCGCG
TCAAGAGAAAGCGCTACAGACCTGCGCCCGGCGGCGGCTCAAGGTCAACAAGAGCTGCGCTGAACCGCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCCAGCCCATACACCTGACGCTGACCGAAGAGCTGCTGCGCACATCGAATGCTGCTGCGCTGCG
GCCAGCCCATACACCTGACGCTGACCGAAGAGCTGCTGCGCACATCGAATGCTGCTGCGCTGCG
X 80 90 100 110 120 130 140

150 160 170 180 190 200
ACGAGGGCCAGACCGGCGATGACCGCTCGCGGCGGCGGCTGAGAGTCCGCTGAGACCGACAT X
ACGAGGGCCAGACCGGCGATGACCGCTCGCGGCGGCGGCTGAGAGTCCGCTGAGACCGACAT X
X 150 160 170 180 190 200

13. US-09-697-123B-11 (1-208)
US-09-697-123B-1 Sequence 17, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.48
Residue Identity = 87% Matches = 181 Mismatches = 27
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCGCTACAGACCTGCGCCCGGCGGCGGCTCAAGGTCAACAAGAGCTGCGCTGAACCGCG
TCAAGAGAAAGCGCTACAGACCTGCGCCCGGCGGCGGCTCAAGGTCAACAAGAGCTGCGCTGAACCGCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCCAGCCCATACACCTGACGCTGACCGAAGAGCTGCTGCGCACATCGAATGCTGCTGCGCTGCG
GCCAGCCCATACACCTGACGCTGACCGAAGAGCTGCTGCGCACATCGAATGCTGCTGCGCTGCG
X 80 90 100 110 120 130 140

Initial Score	-	141	Optimized score	-	167	Significance	-	-0.32
Residue Identity	-	79%	Matches	-	177	Mismatches	-	31
Gaps	-	15	Conservative Substitutions	-			-	0

|||||
TCAAGAGAGAGCCCTACGACCTGCGCCGCGCTACAAAGTCAAGAAAGCTGGGCTGACAGCCG
X 10 20 30 40 50 60 70
GCCAGCCATCACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCAATATGCTGCTGCTG
ATCATCTGATCAACACACGACGCTGCGAGAAAGAGAGCTGCTGCGACCATGATGCTGCTGCTG
80 90 100 110 120 130 140
ACGAGGCC-----CAGACTCGCATGACGCTCGGGCGGTGTCAGAGTCCGCTGAGAGACCG
11
ACGACCCCTCTCAGGCTGCGCCAGCGCCCGCTTATGACTGTCCCGCGGCGTGCAGGTGCGGTAACCG
150 160 170 180 190 200 210
X
ACGACAT
|||||
ACGACAT
220 X

20. US-09-697-123B-11 (1-208)
US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 140 Optimized Score = 168 Significance = -0.34
Residue Identity = 79% Matches = 178 Mismatches = 30
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAGAGCGCTACGACCTGCGCCGCGGTGCGCTACAAAGTCAAGAAAGCTGCGCTGAAACCGG
TCAAGAGAGAGCGCTACGACCTGCGCCGCGGTGCGCTACAAAGTCAAGAAAGCTGCGCTGAAACCGG
X 10 20 30 40 50 60 70
GCCAGCCATCACCACTGACGCTGACCGAGAAAGCTGCTGCGACCATCAATATGCTGCTGCTGCTG
11
GCGACCCCTCTCAGGCTGCGCCAGCGCCCGCTTATGACTGTCCCGCGGCGTGCAGGTGCGGTAACCG
80 90 100 110 120 130 140
ACGAGGCCAG-----CCGCGATGACGCTCGGGCGGTGTCAGAGTCCGCTGAGAGACCG
11
ACGACCCCTCTCAGGCTGCGCCAGCGCCCGCTTATGACTGTCCCGCGGCGTGCAGGTGCGGTAACCG
150 160 170 180 190 200 210
X
ACGACAT
|||||
ACGACAT
220 X

21. US-09-697-123B-11 (1-208)
US-09-697-123B-2 Sequence 21, Application US/09697123B

Initial Score = 138 Optimized Score = 169 Significance = -0.38
Residue Identity = 82% Matches = 176 Mismatches = 32
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAGAGCGCTACGACCTGCGCCGCGGTGCGCTACAAAGTCAAGAAAGCTGCGCTGAAACCGG
TCAAGAGAGAGCGCTACGACCTGCGCCGCGGTGCGCTACAAAGTCAAGAAAGCTGCGCTGAAACCGG
X 10 20 30 40 50 60 70
GCCAGCC-----CATCACCAGCTGACGCTGACCGAGAAAGCTGCTGCGACCATCAATATGCTGCTGCTG
11
CCAATCGGCTCAGGTACGACCAACGACCTTCAACCGAGAAAGAGCTGCTGCGACCATCAATATGCTGCTGCTG
80 90 100 110 120 130 140

140 150 160 170 180 190 200
GCTTGCACAGAGGCCAGACCGCGATGACCGCTCGGGCGGTGAGAGTCCGGTGCAGACCGACGAT
|||||
GCTTGCACAGAGGCCAGACCGCGATGACCGCTCGGGCGGTGAGAGTCCGGTGCAGACCGACGAT
150 160 170 180 190 200 210 X

22. US-09-697-123B-11 (1-208)
US-09-697-123B-1 Sequence 14, Application US/09697123B

Initial Score = 129 Optimized Score = 166 Significance = -0.56
Residue Identity = 80% Matches = 173 Mismatches = 35
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAGAGCGCTACGACCTGCGCCGCGGTGCGCTACAAAGTCAAGAAAGCTGCGCTGAAACCGG
TCAAGAGAGAGCGCTACGACCTGCGCCAGCGCTGCGCTACAAAGTCAAGAAAGCTGCGCTGAAACCGG
X 10 20 30 40 50 60 70
GCCAGCC-----CATCACCAGCTGACGCTGACCGAGAAAGCTGCTGCGACCATCAATATGCTGCTGCTGCTG
11
CCGAGTGGCGCGCTACCGCGCTTCAACACGCTGACCGAGCGGATGCTGCTGCGACCATCAATATGCTGCTGCTGCTG
80 90 100 110 120 130 140
140 150 160 170 180 190 200
GCTTGCACAGAGGCCAGACCGCGATGACCGCTCGGGCGGTGTCAGAGTCCGCTGAGAGACCGACGAT
11
GCCGTCAGAGAGGCCAGACCGCGATGACCGCTCGGGCGGTGTCAGAGTCCGCTGAGAGACCGACGAT
150 160 170 180 190 200 210 X

23. US-09-697-123B-11 (1-208)
US-09-697-123B-1 Sequence 18, Application US/09697123B

Initial Score = 121 Optimized Score = 166 Significance = -0.72
Residue Identity = 81% Matches = 172 Mismatches = 36
Gaps = 3 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAGAGCGCTACGACCTGCGCCGCGGTGCGCTACAAAGTCAAGAAAGCTGCGCTGAAACCGG
TCAAGAGAGAGCGCTACGACCTGCGCCGCGGTGCGCTACAAAGTCAAGAAAGCTGCGCTGAAACCGG
X 10 20 30 40 50 60 70
GCCAGCCATCACCACTGCG---ACGCTGACCGAGAGAGAGCTGCTGCGACCATCAATATGCTGCTGCTGCTGCTG
11
ACCGGCGCGACGACGACCTGCGACCGCTGACCGAGAGAGAGCTGCTGCGACCATCAATATGCTGCTGCTGCTGCTG
80 90 100 110 120 130 140
150 160 170 180 190 200
TGCAGAGAGAGCGCCAGACCGCGATGACCGCTCGGGCGGTGCGCTGCGAGAGAGAGCTGCGCTGAAACCGG
11
TGCATCAGAGAGCGCCAG
150 160 170 180 190 200 210

24. US-09-697-123B-11 (1-208)
US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score = 108 Optimized Score = 168 Significance = -0.98
Residue Identity = 81% Matches = 175 Mismatches = 33
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAGAGCGCTACGACCTGCGCCGCGGTGCGCTACAAAGTCAAGAAAGCTGCGCTGAAACCGG
TCAAGAGAGAGCGCTACGACCTGCGCCGCGGTGCGCTACAAAGTCAAGAAAGCTGCGCTGAAACCGG
X 10 20 30 40 50 60 70
GCCAGCC-----CATCACCAGCTGACGCTGACCGAGAAAGCTGCTGCGACCATCAATATGCTGCTGCTGCTGCTG
11
CCAATCGGCTCAGGTACGACCAACGACCTTCAACCGAGAAAGAGCTGCTGCGACCATCAATATGCTGCTGCTGCTGCTG
80 90 100 110 120 130 140

```

      80      90      100      110      120      130
GCCAGCC-----CATCACCAGCTGACCGAGAGAGCGTCCGCCACCATCGAATACCTGTC
      |||      |||      |||      |||      |||      |||
      80      90      100      110      120      130
CCAACCGGCTGTGTGACTGCCACACGCTCACCAGAGAGACGTGCGCACCATCGGGTACTGTGC
      |||      |||      |||      |||      |||      |||
      140      150      160      170      180      190      200
GCTTGCACGAGGGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGGTGGTCGAGACCGAGCAT
      |||      |||      |||      |||      |||      |||
      140      150      160      170      180      190      200
GCTTGCACGAGGGCCAGACCGCGATGACCGCTCCGGCGGTGTCGAGGTGGTCGAGCAT
      |||      |||      |||      |||      |||      |||
      150      160      170      180      190      200      210
X

```

25. US-09-697-123B-11 (1-208)

US-09-697-123B-2 Sequence 25, Application US/09697123B

Initial Score = 19 Optimized Score = 19 Significance = -2.75
 Residue Identity = 100% Matches = 19 Mismatches = 0
 Caps = 0 Conservative Substitutions = 0

```

X      10      X
TCAAGGAGAAAGCGCTACGA
|||||
TCAAGGAGAAAGCGCTACGA
X      10      X

```

> 0 <
01 10 Intelligenetics
> 0 <
FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-697-123b-10.res made by shanley on Wed 13 Nov 102 14:05:42-PST.

Query sequence being compared: US-09-697-123b-10 (1-208)
Number of sequences searched: 26
Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123b-10 (1-208) with:
File : US09697123b.seq

100-
N -
U -50-
M -
B -
E -
R -
O -
F 10-
S -
S 5-
E -
O -
U -
E -
C -
E -
S 0-
SCORE 0 23 46 69 92 116 139 162 185 208
STDDEV -2 -1 0 1

PARAMETERS
Similarity matrix Unitary K-tuple
Mismatch penalty 1 Joining penalty 4
Gap penalty 5.00 Window size 30
Gap size penalty 0.33 207
Cutoff score 12
Randomization group 0

SEARCH STATISTICS

Scores: Mean Median Standard Deviation
156 180 49.86
Times: CPU
00:00:00.00 Total Elapsed
00:00:00.00

Number of residues: 5077
Number of sequences searched: 26
Number of scores above cutoff: 25

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.	Length	Score	Sig. Frame
---------------	-------------	------------	--------	-------	------------

1. US-09-697-123b-1 Sequence 10, Application 208 208 208 1.04 0
The list of other best scores is:

Sequence Name	Description	Length	Score	Sig. Frame
2. US-09-697-123b-1 Sequence 11, Application		208	205	0.98 0
3. US-09-697-123b-1 Sequence 15, Application		208	188	0.64 0
4. US-09-697-123b-1 Sequence 19, Application		208	186	0.60 0
5. US-09-697-123b-1 Sequence 16, Application		208	186	0.60 0
6. US-09-697-123b-6 Sequence 6, Application		208	186	0.60 0
7. US-09-697-123b-5 Sequence 5, Application		208	186	0.60 0
8. US-09-697-123b-2 Sequence 2, Application		208	186	0.60 0
9. US-09-697-123b-1 Sequence 1, Application		208	186	0.60 0
10. US-09-697-123b-2 Sequence 22, Application		208	185	0.58 0
11. US-09-697-123b-3 Sequence 3, Application		208	183	0.54 0
12. US-09-697-123b-8 Sequence 8, Application		208	182	0.52 0
13. US-09-697-123b-2 Sequence 23, Application		208	179	0.46 0
14. US-09-697-123b-1 Sequence 17, Application		208	179	0.46 0
15. US-09-697-123b-2 Sequence 24, Application		208	174	0.36 0
16. US-09-697-123b-2 Sequence 20, Application		208	165	0.18 0
17. US-09-697-123b-1 Sequence 12, Application		207	160	0.08 0
18. US-09-697-123b-4 Sequence 4, Application		207	149	0.14 0
19. US-09-697-123b-1 Sequence 13, Application		223	140	0.32 0
20. US-09-697-123b-2 Sequence 21, Application		214	139	0.34 0
21. US-09-697-123b-9 Sequence 9, Application		223	139	0.34 0
22. US-09-697-123b-1 Sequence 14, Application		214	126	0.60 0
23. US-09-697-123b-1 Sequence 18, Application		211	122	0.68 0
24. US-09-697-123b-7 Sequence 7, Application		214	107	0.98 0
25. US-09-697-123b-2 Sequence 25, Application		19	19	-2.75 0

1. US-09-697-123b-10 (1-208)
US-09-697-123b-1 Sequence 10, Application US/09697123b

Initial Score	-	208	Optimized Score	-	208	Significance	-	1.04
Residue Identity	-	100%	Matches	-	208	Mismatches	-	0
Gaps	-	0	Conservative Substitutions	-			-	0

X	10	20	30	40	50	60	70
TCAGAGAGAGCGCTACGACCTGCTGCGTGGCGTGCAGTCAAGTCAAGAGCGCTGACCGCG							
TCAGAGAGAGAGCGCTACGACCTGCTGCGTGGCGTGCAGTCAAGTCAAGAGCGCTGACCGCG							
X	10	20	30	40	50	60	70

GCCAGGCCATTCACCACTGACGCTGACCGTGCAGAGAGAGTGCAGTGCAGTGCAGTGCAGTGC							
GCCAGGCCATTCACCACTGACGCTGACCGTGCAGAGAGAGTGCAGTGCAGTGCAGTGCAGTGC							
GCCAGGCCATTCACCACTGACGCTGACCGTGCAGAGAGAGTGCAGTGCAGTGCAGTGCAGTGC							
X	80	90	100	110	120	130	140

ACGAGGCCATTCACCACTGACGCTGACCGTGCAGAGAGAGTGCAGTGCAGTGCAGTGCAGTGC							
ACGAGGCCATTCACCACTGACGCTGACCGTGCAGAGAGAGTGCAGTGCAGTGCAGTGCAGTGC							
ACGAGGCCATTCACCACTGACGCTGACCGTGCAGAGAGAGTGCAGTGCAGTGCAGTGCAGTGC							
X	150	160	170	180	190	200	X

2. US-09-697-123b-10 (1-208)
US-09-697-123b-1 Sequence 11, Application US/09697123b

Initial Score	-	205	Optimized Score	-	205	Significance	-	0.98
Residue Identity	-	98%	Matches	-	205	Mismatches	-	3
Gaps	-	0	Conservative Substitutions	-			-	0

X	10	20	30	40	50	60	70
TCAGAGAGAGCGCTACGACCTGCTGCGTGGCGTGCAGTCAAGTCAAGAGCGCTGACCGCG							
TCAGAGAGAGCGCTACGACCTGCTGCGTGGCGTGCAGTCAAGTCAAGAGCGCTGACCGCG							
X	10	20	30	40	50	60	70

80 90 100 110 120 130 140
GCCAGCCCATACCAAGCTCGAAGCCTGACCAGGAAGACGTCTGCACCATCATGATTAATTACTTGTCCTGGTCTG
80 90 100 110 120 130 140
GCCAGCCCATACCAAGCTCGAAGCCTGACCAGGAAGACGTCTGCACCATCATGATTAATTACTTGTCCTGGTCTG

150 160 170 180 190 200 X
 ACGAGGGCCAGACCCGATGACCGCTCCGGCGCGTTCAGGTCGCCGTCGACACCGACGACAT
 150 160 170 180 190 200 X
 ACGAGGGCCAGACCCGATGACCGCTCCGGCGCGTTCAGGTCGCCGTCGACACCGACGACAT

3. US-09-697-123B-10 (1-208)
US-09-697-123B-1 Sequence 15, Application US/09697123B

Initial Score	=	188	Optimized Score	=	188	Significance	=	0.64
Residue Identity	=	90%	Matches	=	188	Mismatches	=	20
Gaps	=	0	Conservative Substitutions	=			=	0

X 10 20 30 40 50 60 70
 TCAAGAGAGCCCTACGACCGGCGGCGGTGGGTACAAAGTCAACAAAGCTGGCTGACCCCG
 |||||
 TCAAGAGAGCCCTACGACCGGCGGCGGTGGGTACAAAGTCAACAAAGCTGGCTGACCCCG
 X 10 20 30 40 50 60 70

GCCAGGCCATCACACGCTGCAGCGTGACCAGGAAGACGTGTCGCCACCATGAAATACCTGGTCCGGCTTGC
80 90 100 110 120 130 140
| | | | | | | | | | | | | | | | | |
GTAGCCGATCACACGCTGCAGCGTGACCAGGAAGACGTGTCGCCACCATGCAACTGGTGGCCTGC
80 90 100 110 120 130 140

	150	160	170	180	190	200	
ACGAGGGCGACCGGGATGACCGGCGCGGGGGTGTGTGAGAGTCCCGGTGAGACCGACGACAT	150	160	170	180	190	200	X
ACGAGGGTACGCGCCAGATGACCGTCCCGGGCGCATGAGAGTCCCGGTGAGAACCGACGACAT	150	160	170	180	190	200	X

4. US-09-697-123B-10 (1-208)
US-09-697-123B-1 Sequence 19, Application US/09697123B

Initial Score	=	186	Optimized Score	=	186	Significance	=	0.60
Residue Identity	=	89%	Matches	=	186	Mismatches	=	22
Gaps	=	0	Conservative Substitutions	=			=	0

[illegible][illegible]

150	160	170	180	190	200	X
GGAGGGCCAGACCGGCATGACCGCTCGGGCGGTGTGAGGTGCCGGTCCAGACACGACGACAT						
150	160	170	180	190	200	X
GAGGGCCAGACCGCATGACCGTCCCGGCGGCTGAGTCCCGGTCCAGGTGAGCAGCAT						

5. US-09-697-123B-10 (1-208)
US-09-697-123B-1 Sequence 16, Application US/09697123E

Initial Score =	186	Optimized Score =	186	Significance =	0.60
Residue Identity =	89%	Matches =	186	Mismatches =	22
Gaps =	0	Conservative Substitutions			0

x	10	20	30	40	50	60	70
---	----	----	----	----	----	----	----

TCAAGGAGAAAGCCCTACGACCTGGCGTCGCTGGGTACAAGGTCACAAAGAGCTGGCGCTGAACGCCG
 TCAAGGAGAAAGCCCTACGAGCTGGCGCGCGGTCGCTATAGGTCAACAAAGAGCTGGCGCTGCATGTGCG
 X 10 20 30 40 50 60 70

GCCAGCCCATCACCACTCTGGACGCTGACGAGAACGCTCTGCCACATTCGAATACCTGCTCGGTGG
 |||||
 GCGAGCCCATCACAGTGTGACGCTGACGAGAACGCTCTGCCACATTCGAATATCTGCTCGGTGG
 |||||
 80 90 100 110 120 130 140

ACGAGGCCACGACCGCGGATGACCGCTCCGGGGCGGTGGAGGTGCGGTCACGACCCACACAT
150 160 170 180 190 200 X
ACGAGGCTGACACCGAGTACCGTCCGGCGCGGTGAGGTGCGGTTGAAACCCACACAT
150 160 170 180 190 200 X

6. US-09-697-123B-10 (1-208)
US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score	-	186	Optimized Score	-	186	Significance	-	0.60
Residue Identity	-	89%	Matches	-	186	Mismatches	-	22
Gaps	-	0	Conservative Substitutions	-			-	0

X	10	20	30	40	50	60	70
TCAAGAGAAGCGCTACGACCTGCTGCGCTGGTGGTACAAAGTCAACAAGAGCTGGCGCTGACGCCGCG							
TCAAGGAGAGCGCTACACCTGGCCCGCTCGTGCATATAGTCAACAAGAGCTGGCGCTGCATCTGCG							
X	10	20	30	40	50	60	70

GCCAGCCCCATCACCAGCTGCAGCGCTGACCGAGAGACGCTGTCGCCACCATGTAATACCTGGTCCGGCTTGC
 80 90 100 110 120 130 140
 GCGAGGCCATTCACGTCGTGTCGACGCTGACCGAAGAGAGCTGTGGCCCATCGAATATCTCTGTCGCTTGC
 80 90 100 110 120 130 140

ACAGAGGGCCAGACCGCGATGACCGTCCGGGGGGGTGTCGAGTCCGGTCGAGACCCAGACAT
 150 160 170 180 190 200 X
 |||||
 ACAGAGGTGACACACAGATGACCGTCCGGGGGGGTGTCGAGTCCGGTCGAGACCCAGACAT
 150 160 170 180 190 200 X

7. US-09-697-123B-10 (1-208)
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial score	-	186	Optimized Score	-	186	Significance	-	0.60
Residue Identity	-	89%	Matches	-	186	Mismatches	-	22
Gaps	-	0	Conservative Substitutions	-			-	0

X 10 20 30 40 50 60 70
 TCAAGGAAACCGTACGACCTGGCTGGGTGGGTCAAGAAGTCAAGAGTGGGCTGACGCCG
 X 10 20 30 40 50 60 70
 TCAAGGAAACCGTACGACCTGGCTGGGTGGGTCAAGAAGTCAAGAGTGGGCTGACGCCG

80 90 100 110 120 130 140
 GCGACCCATTCACAGTTCGACCGTCACCGAGAGAGCGTGTGCCACATCGAATACCTGGTCCGCTTCC
 80 90 100 110 120 130 140
 GCGACCCATTCACGTCGTTCGACCGTCACCGAGAGAGCGTGTGCCACATCGAATATCTGGTCCGCTTCC

150	160	170	180	190	200	X
ACGAGGCGCCACAGACCGCGATGACCGCTCCGGCGGCGGTGTTCGAGAGTCCCGGTGAGAACCGACACAT						
CGAGGCGTCACAGCACCATGACCGTCCGGCGGGCGGCGGTTCGAGAGTCCCGGTGAGAACCGACACAT						
150	160	170	180	190	200	X

8. US-09-697-123B-10 (1-208)
US-09-697-123B-2 Sequence 2, Application US/09697123B

[illegible]

14. US-09-697-123B-10 (1-208)
US-09-697-123B-1 Sequence 17, Application US/09697123B

mutual score	=	1/9	Optimized score	=	1/9	Significance	=	0.46
Residue Identity	=	86%	Matches	=	179	Mismatches	=	29
Gaps	=	0	Conservative Substitutions	=	0			

X	10	20	30	40	50	60	70
TCAAGGAGAGCCCTACGACCTGGCTCGCGTGGGTGGGTACAAGGTCAACAAAGAAGCTCGGCTGACCC							
X	10	20	30	40	50	60	70
TCAAGGAGAGCGCTACGACCTCGCGGGGGTGGCGGTACAAGGTCAACAAAGAAGCTCGGCTGACCC							

80 90 100 110 120 130 140
 GCCAGCCATCAACAGCTGACGGTGCAGAGAACCTCTCCACACATCGAATACCTGTCGGCTTG
 |||||
 CGTCCCGATCAAGACACCTCTGACCGAAGAGAGCGTCTCCACCATCGAGTACCTGTCGGCTTG
 80 90 100 110 120 130 140

150	160	170	180	190	200	X
ACGAGGGCCAGACCCGCGATTCGCGGGGGTGAGAGTGGCCGTCGAGACCGACGACAT						
ACGAGGGCCACACACGATTCGCCGGGGGAGTGCAGAGTGGCGGTTGAAACCCACGACAT						
150	160	170	180	190	200	X

15. US-09-697-123B-10 (1-208)
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score	=	174	Optimized Score	=	174	Significance	=	0.36
Residue Identity	=	83%	Matches	=	174	Mismatches	=	34
Gaps	=	0	Conservative Substitutions	=	0			

X
10 20 30 40 50 60 70
TCAAGGAGAAAGCGTACGACGCTGCGGTGGGTGCGTACAAAGTACACAGAAGTGGGCTGAACGG
TCAAGGAGAAAGCGTACGACGCTGCGGTGGGTGCGTACAAAGTACACAGAAGTGGGCTGAACGG
TCAAGGAGAAAGCGTACGACGCTGCGGTGGGTGCGTACAAAGTACACAGAAGTGGGCTGAACGG
X
10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCCAGCCCTTCACCGATGACGTGACCGAGAAGAGTGTGCACCATGGATTACTGTCCGCTTT
| | | | |
AGAAATGCCCAACCAACGACCCTTGACCGAAGGAGCTGTGCGCACCATGGATTACTGTGGCTTT
80 90 100 110 120 130 140

ACGAGGGCCAGACCCTGCATGACCGTCCGGCGGTGTGAGGTGCCGTTCAGACCCACGACAT
X
150 160 170 180 190 200

ACGAGGGCCAGACCCTGCATGAGGTCCGGCGGTGTGAGGTGCCGTTCAGACCCACGACAT
X
150 160 170 180 190 200

16. US-09-697-123B-10 (1-208)
US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score	=	165	Optimized Score	=	176	Significance	=	0.18
Residue Identity	=	87%	Matches	=	182	Mismatches	=	23
Gaps	=	3	Conservative Substitutions	=			=	0

[illegible][illegible]

150	160	170	180	190	200	X
GAGGGCCAGACCGCGATGACGCGTCCGGGCGGTTCGAGGTGCCGGTTCAGAGCCAGCAGCAT						
150	160	170	180	190	200	X
GAGGGCCAGCCACCATGACGCGTCCC---CGCATTCGAGGTGCCGGTTCGAGAGCCAGCAGCAT						

17. US-09-697-123B-10 (1-208)
US-09-697-123B-1 Sequence 12, Application US/09697123B

Initial Score	-	160	Optimized Score	-	179	Significance	-	0.08
Residue Identity	-	88%	Matches	-	184	Mismatches	-	23
Gaps	-	1	Conservative Substitutions	-	0		-	0

X 10 20 30 40 50 60 70
 TCAAGGAAAGCCCTACGACCTGCGCTCGGTCGTCAAGTCAACAAGAAAGCTCGGCTCAAGCGCC
 X 10 20 30 40 50 60 70
 TCAAGGAAAGCCCTACGACCTGCGCTCGGTCGTCAAGTCAACAAGAAAGCTCGGCTCAAGCGCC

[illegible]

150	160	170	180	190	200	X
GAGGGCCGACCGGGAGTACCGCTCCGGGGGGGTGTCAAGTGGCGGTGAGAGCCGACGACAT						
150	160	170	180	190	200	X
GGGGCCGACCGAGATGACCGTCCGGGGGACCCAGTAGTGCCGTGGACACCGACGACAT						

18. US-09-697-123B-10 (1-208)
US-09-697-123B-4 Sequence 4, Application US/09697123E

Residue Identity	=	86%	Matched score	=	-0.14	Significance
Gaps	=	1	Optimized	=	1/5	
			Matches	=	180	Mismatches
			Conservative Substitutions	=	27	

X 10 20 30 40 50 60 70
TCAAGGAAAGCCCTACGACTCGGCTCGGCTCGTCAAGTCAACAGAAAGCTCGGCTGAACCCCTC
X 10 20 30 40 50 60 70
TCAAGGAAAGCCCTACGACTCGGCTCGGCTCGTCAAGTCAACAGAAAGCTCGGCTGAACCCCTC

[illegible]

150	160	170	180	190	200	X
ACGAGGGCCAGACCGCGATGACGGCTCCGGCGGTTGAGAGTGGCGGTGAGAGCCACGACAT						
150	160	170	180	190	200	X
ACGGGGCTCAGACACAGATGA-CGTTCCGGGGGGAGACCGAGTTCCGTTGGAGACCGACACAT						
150	160	170	180	190	200	X

19. US-09-697-123B-10 (1-208)
US-09-697-123B-1 Sequence 13, Application US/09697123B

Residue score	-	140	Optimized score	-	166	Significance	-	-0.32
Identity	-	78%	Matches	-	176	Mismatches	-	32
Gaps	-	15	Conservative Substitutions	-	0		-	0

X
10 20 30 40 50 60 70
TCAAGAGAGCCCTACGACCTGGCTGGCTGGCTACAGGCTCAACAGAGCTGGGCTGAACGCCG

X	10	20	30	40	50	60	70
TCAAGGAGACCGCTACGACTTGCGTGGGTGGTACAAGTCAACAAGAAGCTGGCCTGAACGCCG							
TCAAAGAAACCGCTACGACTTGCCCGGTGGCGGTGATCAAGGAAGTAAGAGCTTTGGCGGGG							
X	10	20	30	40	50	60	70

> 0 <
01 10 Intelligenetics
> 0 <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-697-123b-9.res made by shanley on Wed 13 Nov 102 14:05:06-PST.

Query sequence being compared: US-09-697-123B-9 (1-223)
Number of sequences searched: 26
Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123B-9 (1-223) with:
File : US09697123B.seq

```

100-
-
N -
U -
M -
B -
E -
R -
O -
F 10-
S -
E 5-
O -
U -
N -
C -
E -
S 0-
SCORE 0 125 50 1 74 99 1 124 149 1 173 198 223
STDEV -2 -1 0 0 0 1 1 1 1 1 1 1 1
PARAMETERS
Similarity matrix Unitary 1 K-tuple 4
Mismatch penalty 5.00 Joining penalty 30
Gap penalty 0.33 Window size 207
Gap size penalty 13
Cutoff score 13
Randomization group 0
SEARCH STATISTICS
Scores: Mean 127 Median 139 Standard Deviation 44.15
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
Number of residues: 5077
Number of sequences searched: 26
Number of scores above cutoff: 25

```

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.	Length	Score	Score	Sig. Frame
US-09-697-123B-9	Sequence 9, Application U	223	223	223	2.17	0

1. US-09-697-123B-9 Sequence 9, Application U 223 223 223 2.17 0
The list of other best scores is:

Sequence Name	Description	Length	Score	Score	Sig. Frame
2. US-09-697-123B-1	Sequence 13, Application 13	223	195	195	1.54 0
3. US-09-697-123B-1	Sequence 15, Application 15	208	148	175	0.48 0
4. US-09-697-123B-1	Sequence 19, Application 19	208	146	173	0.43 0
5. US-09-697-123B-1	Sequence 1, Application 1	208	146	173	0.43 0
6. US-09-697-123B-1	Sequence 16, Application 16	208	144	170	0.39 0
7. US-09-697-123B-6	Sequence 6, Application 6	208	144	170	0.39 0
8. US-09-697-123B-5	Sequence 5, Application 5	208	144	170	0.39 0
9. US-09-697-123B-2	Sequence 2, Application 2	208	144	170	0.39 0
10. US-09-697-123B-2	Sequence 22, Application 22	208	143	169	0.39 0
11. US-09-697-123B-1	Sequence 11, Application 11	208	140	168	0.29 0
12. US-09-697-123B-1	Sequence 10, Application 10	208	139	167	0.27 0
13. US-09-697-123B-2	Sequence 23, Application 23	208	138	163	0.25 0
14. US-09-697-123B-8	Sequence 8, Application 8	208	138	165	0.25 0
15. US-09-697-123B-1	Sequence 17, Application 17	208	137	164	0.23 0
16. US-09-697-123B-2	Sequence 24, Application 24	208	136	160	0.20 0
17. US-09-697-123B-3	Sequence 3, Application 3	208	133	171	0.14 0
18. US-09-697-123B-2	Sequence 20, Application 20	205	131	163	0.09 0
19. US-09-697-123B-4	Sequence 4, Application 4	207	127	163	0.00 0
20. US-09-697-123B-1	Sequence 18, Application 18	211	116	150	-0.25 0
21. US-09-697-123B-1	Sequence 12, Application 12	207	112	154	-0.34 0
22. US-09-697-123B-7	Sequence 7, Application 7	214	91	154	-0.82 0
23. US-09-697-123B-2	Sequence 21, Application 21	214	83	152	-1.00 0
24. US-09-697-123B-1	Sequence 14, Application 14	214	81	156	-1.04 0
25. US-09-697-123B-2	Sequence 25, Application 25	19	19	19	-2.45 0

1. US-09-697-123B-9 (1-223)
US-09-697-123B-9 Sequence 9, Application US/09697123B
Initial Score = 223 Optimized Score = 223 Significance = 2.17
Residue Identity = 100% Matches = 223 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

```

X 10 20 30 40 50 60 70
TCAAGAGAGAGCGCTACGACGCTGCGCCCGCTGCGCCGCTACAAAGCTCAAGAGAGAGCGCGCG
TCAAGAGAGAGCGCTACGACGCTGCGCCCGCTGCGCCCGCTACAAAGCTCAAGAGAGAGCGCGCG
X 10 20 30 40 50 60 70
GCGAGCCGATCAGCTGTCACGCTGACGCGTGAAGAGAGAGCTGCGCCGCTGCGCCGCTGCGCGCG
GCGAGCCGATCAGCTGTCACGCTGACGCGTGAAGAGAGAGCTGCGCCGCTGCGCCGCTGCGCGCG
GCGAGCCGATCAGCTGTCACGCTGACGCGTGAAGAGAGAGCTGCGCCGCTGCGCCGCTGCGCGCG
X 10 20 30 40 50 60 70
ACCAAGCCCGGATGAGATGCGCGCGCGCGCTGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
ACCAAGCCCGGATGAGATGCGCGCGCGCGCTGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
X 10 20 30 40 50 60 70
ACCAAGCCCGGATGAGATGCGCGCGCGCGCTGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
ACCAAGCCCGGATGAGATGCGCGCGCGCGCTGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
X 10 20 30 40 50 60 70
ACGACAT 11111111
ACGACAT 220 X

```

2. US-09-697-123B-9 (1-223)
US-09-697-123B-1 Sequence 13, Application US/09697123B
Initial Score = 195 Optimized Score = 195 Significance = 1.54

X 10 20 30 40 50 60 70
 TCAAGGAGAGCGGTACGACCTGGCCCGTCCGCGGTACAAGGCAACGAAGAGCTGTGTCACGGC
 X 10 20 30 40 50 60 70
 TCAAGGAGAGCGGTACGACCTGGCCCGTCCGCGGTACAAGGCAACGAAGAGCTGTGTCACGGC

ACCAAGCCCGTACGGATGGCCAGCCCCCGTACGTACTGTCCCGCGGCGATCGAGAGTCCGGTGGAGACCG
150 160 170 180 190 200 210

TCAGAGAGAGCCCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGACTGGTCTGCACGCC
|||||
TCAGAGAGAGCCCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGACTGGTCTGCACGCC
X 10 20 30 40 50 60 70
GCGAGCCGATCAGCTGCTCCAGCCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGCAGCTGC
|||||
GCGAGCCGATCAGCTGCTCCAGCCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGCAGCTGC
80 90 100 110 120 130 140
ACCAAGCCGCTACGAGATGGCCAGCCCGCGCTACATGACTGTCCCGGCGCATCGAGATGCCGTGGAGACCG
|||||
ACGAGGCCGAGA-----CCGCGATGACCCGCTCGGCGGTGTCAGAGTCCGCGTGGAGACCG
150 160 170 180 190 200
220 X
ACGACAT
|||||
ACGACAT
X
12. US-09-697-123B-9 (1-223)
US-09-697-123B-1 Sequence 10, Application US/09697123B
Initial Score = 139 Optimized Score = 167 Significance = 0.27
Residue Identity = 79% Matches = 177 Mismatches = 31
Gaps = 15 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGAGCTGGTCTGCACGCC
|||||
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGAGCTGGTCTGCACGCC
X 10 20 30 40 50 60 70
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGAGCTGGTCTGCACGCC
80 90 100 110 120 130 140
GCGAGCCGATCAGCTGCTCCAGCCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGCAGCTGC
|||||
GCGAGCCGATCAGCTGCTCCAGCCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGCAGCTGC
150 160 170 180 190 200
ACCAAGCCGCTACGAGATGGCCAGCCCGCGCTACATGACTGTCCCGGCGCATCGAGATGCCGTGGAGACCG
|||||
ACGAGGCCGAGA-----CCGCGATGACCCGCTCGGCGGTGTCAGAGTCCGCGTGGAGACCG
150 160 170 180 190 200
220 X
ACGACAT
|||||
ACGACAT
X
13. US-09-697-123B-9 (1-223)
US-09-697-123B-2 Sequence 23, Application US/09697123B
Initial Score = 138 Optimized Score = 163 Significance = 0.25
Residue Identity = 77% Matches = 173 Mismatches = 35
Gaps = 15 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGAGCTGGTCTGCACGCC
|||||
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGAGCTGGTCTGCACGCC
X 10 20 30 40 50 60 70
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGAGCTGGTCTGCACGCC
80 90 100 110 120 130 140
GCGAGCCGATCAGCTGCTCCAGCCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGCAGCTGC
|||||
GCGAGCCGATCAGCTGCTCCAGCCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGCAGCTGC
150 160 170 180 190 200
ACCAAGCCGCTACGAGATGGCCAGCCCGCGCTACATGACTGTCCCGGCGCATCGAGATGCCGTGGAGACCG
|||||
ACGAGGCCGAGA-----CCGCGATGACCCGCTCGGCGGTGTCAGAGTCCGCGTGGAGACCG
150 160 170 180 190 200
220 X
ACGACAT
|||||
ACGACAT
X

80 90 100 110 120 130 140
ACCAAGCCGCTACGAGATGGCCAGCCCGCGCTACATGACTGTCCCGGCGCATCGAGATGCCGTGGAGACCG
|||||
ACGAGGCCGAGA-----CCGCGATGACCCGCTCGGCGGTGTCAGAGTCCGCGTGGAGACCG
150 160 170 180 190 200
220 X
ACGACAT
|||||
ACGACAT
X
14. US-09-697-123B-8 (1-223)
US-09-697-123B-8 Sequence 8, Application US/09697123B
Initial Score = 138 Optimized Score = 165 Significance = 0.25
Residue Identity = 78% Matches = 175 Mismatches = 33
Gaps = 15 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGAGCTGGTCTGCACGCC
|||||
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGAGCTGGTCTGCACGCC
X 10 20 30 40 50 60 70
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGAGCTGGTCTGCACGCC
80 90 100 110 120 130 140
GCGAGCCGATCAGCTGCTCCAGCCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGCAGCTGC
|||||
GCGAGCCGATCAGCTGCTCCAGCCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGCAGCTGC
150 160 170 180 190 200
ACCAAGCCGCTACGAGATGGCCAGCCCGCGCTACATGACTGTCCCGGCGCATCGAGATGCCGTGGAGACCG
|||||
ACGAGGCCGAGA-----CCGCGATGACCCGCTCGGCGGTGTCAGAGTCCGCGTGGAGACCG
150 160 170 180 190 200
220 X
ACGACAT
|||||
ACGACAT
X
15. US-09-697-123B-9 (1-223)
US-09-697-123B-1 Sequence 17, Application US/09697123B
Initial Score = 137 Optimized Score = 164 Significance = 0.23
Residue Identity = 78% Matches = 174 Mismatches = 34
Gaps = 15 Conservative Substitutions = 0
X 10 20 30 40 50 60 70
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGAGCTGGTCTGCACGCC
|||||
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGAGCTGGTCTGCACGCC
X 10 20 30 40 50 60 70
TCAGAGAGAGCCGCTACGACCTGCGCCGCGCTACAAAGTCAACAGAGAGCTGGTCTGCACGCC
80 90 100 110 120 130 140
GCGAGCCGATCAGCTGCTCCAGCCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGCAGCTGC
|||||
GCGAGCCGATCAGCTGCTCCAGCCTACGAGAGAGAGCTGCGGACCATGAAATACCTGGTGCAGCTGC
150 160 170 180 190 200
ACCAAGCCGCTACGAGATGGCCAGCCCGCGCTACATGACTGTCCCGGCGCATCGAGATGCCGTGGAGACCG
|||||
ACGAGGCCGAGA-----CCGCGATGACCCGCTCGGCGGTGTCAGAGTCCGCGTGGAGACCG
150 160 170 180 190 200
220 X
ACGACAT
|||||
ACGACAT
X

220 X
AGACCGACGACAT
|||||
AGACCGACGACAT
210 X

25. US-09-697-123B-9 (1-223)
US-09-697-123B-2 Sequence 25, Application US/09697123B
Initial Score = 19 Optimized Score = 19 Significance = -2.45
Residue Identity = 100% Matches = 19 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0
X 10 X
TCAAGGAGAAAGCGCTACGA
|||||
TCAAGGAGAAAGCGCTACGA
X 10 X

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Query sequence being compared:	US-09-697-123B-8 (1-208)
Number of sequences searched:	26
Number of scores above cutoff:	25

Results of the initial comparison of US-09-697-123B-8 (1-208) with:
File : US09697123B.seq

100-
-
N
-
U
50-
-
M
-
B
-
E
-
R
-
O
10-
F
-
S
5-
E
O
U
-
N
C
E
-
S
C
E
-
0-
SCORE
0
STDEV
-2
-1
0
1

PARAMETERS			
Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	5.00	Window size	208
Gap size penalty	0.33		
Cutoff score	12		
Randomization group	0		

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
---------	------	--------	--------------------

Times:	CPU	Total Elapsed
0.000	0.000	0.000

Number of residues:	5077
Number of sequences searched:	26
Number of scores above cutoff:	25

The scores below are sorted by initial score. Significance is calculated based on initial score

A 100% identical sequence to the query sequence was found:

Sequence Name	Description	Int. Opt. Length Score	Score	Sig.	Frame
---------------	-------------	---------------------------	-------	------	-------

1. US-09-697-123B-8 Sequence 8, Application U	208	208	208	1.11	0
---	-----	-----	-----	------	---

Sequence Name	Description	Length	Score	Score	Sig.	Frame
---------------	-------------	--------	-------	-------	------	-------

25.	US-09-697-123B-2	Sequence 25.	Application	19	19	19	-2.77	0
24.	US-09-697-123B-1	Sequence 18.	Application	211	117	164	-0.76	0
23.	US-09-697-123B-2	Sequence 21.	Application	214	122	170	-0.76	0
22.	US-09-697-123B-1	Sequence 14.	Application	214	118	168	-0.74	0
21.	US-09-697-123B-7	Sequence 7.	Application	214	127	164	-0.55	0
20.	US-09-697-123B-4	Sequence 9.	Application	223	138	165	-0.33	0
19.	US-09-697-123B-1	Sequence 12.	Application	207	153	174	-0.02	0
18.	US-09-697-123B-1	Sequence 13.	Application	207	153	174	-0.02	0
17.	US-09-697-123B-2	Sequence 20.	Application	205	163	173	0.18	0
16.	US-09-697-123B-2	Sequence 23.	Application	208	177	177	0.47	0
15.	US-09-697-123B-2	Sequence 2.	Application	208	178	178	0.49	0
14.	US-09-697-123B-2	Sequence 22.	Application	208	179	179	0.51	0
13.	US-09-697-123B-2	Sequence 24.	Application	208	181	181	0.55	0
12.	US-09-697-123B-2	Sequence 22.	Application	208	182	182	0.58	0
11.	US-09-697-123B-5	Sequence 5.	Application	208	182	182	0.58	0
10.	US-09-697-123B-1	Sequence 10.	Application	208	182	182	0.58	0
9.	US-09-697-123B-1	Sequence 16.	Application	208	182	182	0.58	0
8.	US-09-697-123B-1	Sequence 1.	Application	208	183	183	0.60	0
7.	US-09-697-123B-1	Sequence 19.	Application	208	183	183	0.60	0
6.	US-09-697-123B-3	Sequence 3.	Application	208	184	184	0.62	0
5.	US-09-697-123B-1	Sequence 11.	Application	208	184	184	0.62	0
4.	US-09-697-123B-1	Sequence 15.	Application	208	185	185	0.64	0
3.	US-09-697-123B-1	Sequence 17.	Application	208	187	187	0.68	0
2.	US-09-697-123B-1	Sequence 17.	Application	208	187	187	0.68	0
1.	US-09-697-123B-1	Sequence 17.	Application	208	187	187	0.68	0

1. US-09-697-123B-8 (1-208)
US-09-697-123B-8 Sequence 8, Application US/09697123B

Initial Score	-	208	Optimized Score	-	208	Significance	-	1.11
Residue Identity	-	100%	Matches	-	208	Mismatches	-	0
Gaps	-	0	Conservative Substitutions	-			-	0

X 10 20 30 40 50 60 70
TCAAGGAAAGCGCTACGACCTGGCCGTTCCGCCGATCAAGTCAACAAAGAGCTGGGCTGAACGCC
TCAAGGAAAGCGCTACGACCTGGCCGTTCCGCCGATCAAGTCAACAAAGAGCTGGGCTGAACGCC
TCAAGGAAAGCGCTACGACCTGGCCGTTCCGCCGATCAAGTCAACAAAGAGCTGGGCTGAACGCC

ATCATCGATCACCACGACGAGCGTGACCCGACAGACGTCGTGGCACCATCGAGTATCTGTGTCGGCTG
 |||||
 ATCATCGATCACCACGACGAGCGTGACCCGACAGACGTCGTGGCACCATCGAGTATCTGTGTCGGCTG
 80 90 100 110 120 130 140
 ATCATCGATCACCACGACGAGCGTGACCCGACAGACGTCGTGGCACCATCGAGTATCTGTGTCGGCTG
 80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGGCCAGCCACGATGACCGCGGGGGGTCAGTGCCTGGGAACCGACGACAT
ACGGGGCCAGCCACGATGACCGCGGGGGTCAGTGCCTGGGAACCGACGACAT
150 160 170 180 190 200 X

2. US-09-697-123B-8 (1-208)
US-09-697-123B-1 Sequence 17, Application US/09697123B

Initial Score	-	187	Optimized Score	-	187	Significance	-	0.68
Residue Identity	-	898	Matches	-	187	Mismatches	-	211
Gaps	-	0	Conservative Substitutions	-			-	0

X	10	20	30	40	50	60	70
TC	AA	GA	AG	AG	CC	CT	AG
AC	CT	AG	CT	GG	CC	GT	GG
CG	CG	GT	GG	CG	AT	ACA	AG
GT	CG	CG	CG	CG	AT	CA	AG
CA	AG	AG	AG	AG	AG	AG	CT
AG	CT	GG	CC	CT	GG	CC	CT
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
TC	AA	GA	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
CA	AG	AG	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
TC	AA	GA	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
CA	AG	AG	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
TC	AA	GA	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
CA	AG	AG	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
TC	AA	GA	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
CA	AG	AG	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
TC	AA	GA	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
CA	AG	AG	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
TC	AA	GA	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
CA	AG	AG	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
TC	AA	GA	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG	CG	CG
CA	AG	AG	AG	AG	AG	AG	AG
AG	AG	AG	AG	AG	AG	AG	AG
GC	CT	AG	CT	AG	CT	AG	CT
CT	AG	CT	GG	CG	CG</		

x	10	20	30	40	50	60	70
---	----	----	----	----	----	----	----

8. US-09-697-123B-8 (1-208)
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.58
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAACCGGTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGGAACACCA
TCAAGAGAACCGGTACGACCTGGCCCGCTGCGCTGATAGGTCAACAGAAAGCTCGGGCTGCATGTGC
X 10 20 30 40 50 60 70

ATCATCCGATCACCACGACGCTGACCGGAGAAAGAGCTGCGCCATCGATGATGTGCTGCGCTGC
GCGACCCGATCAGCTGCTGCGACGCTGACGAAAGAGCTGCTGCGCCATCGAATATCTGCTGCTTGC
X 80 90 100 110 120 130 140

ACGAGGCCAGGACGATGACGCTGCGCGGCGGAGTGAAGTGAAGAAACCGACACAT
ACGAGGCCAGGACGATGACGCTGCGCGGCGGAGTGAAGTGAAGAAACCGACACAT
X 150 160 170 180 190 200

9. US-09-697-123b-8 (1-208)
US-09-697-123b-1 Sequence 10, Application US/09697123b

Initial Score = 182 Optimized Score = 182 Significance = 0.58
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAACCGGTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGGAACACCA
TCAAGAGAACCGGTACGACCTGGCCCGCTGCGCTGATAGGTCAACAGAAAGCTCGGGCTGAACGCGG
X 10 20 30 40 50 60 70

ATCATCCGATCACCACGACGCTGACCGGAGAAAGAGCTGCGCCATCGATGATGTGCTGCGCTGC
GCGACCCGATCAGCTGCTGCGACGCTGACGAAAGAGCTGCTGCGCCATCGAATATCTGCTGCTTGC
X 80 90 100 110 120 130 140

ACGAGGCCAGGACGATGACGCTGCGCGGCGGAGTGAAGTGAAGAAACCGACACAT
ACGAGGCCAGGACGATGACGCTGCGCGGCGGAGTGAAGTGAAGAAACCGACACAT
X 150 160 170 180 190 200

10. US-09-697-123b-8 (1-208)
US-09-697-123b-6 Sequence 6, Application US/09697123b

Initial Score = 182 Optimized Score = 182 Significance = 0.58
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAACCGGTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGGAACACCA
TCAAGAGAACCGGTACGACCTGGCCCGCTGCGCTGATAGGTCAACAGAAAGCTCGGGCTGCATGTGC
X 10 20 30 40 50 60 70

ATCATCCGATCACCACGACGCTGACCGGAGAAAGAGCTGCGCCATCGATGATGTGCTGCGCTGC
GCGACCCGATCAGCTGCTGCGACGCTGACGAAAGAGCTGCTGCGCCATCGAATATCTGCTGCTTGC
X 80 90 100 110 120 130 140

ACGAGGCCAGGACGATGACGCTGCGCGGCGGAGTGAAGTGAAGAAACCGACACAT
ACGAGGCCAGGACGATGACGCTGCGCGGCGGAGTGAAGTGAAGAAACCGACACAT
X 150 160 170 180 190 200

11. US-09-697-123b-8 (1-208)
US-09-697-123b-5 Sequence 5, Application US/09697123b

Initial Score = 182 Optimized Score = 182 Significance = 0.58
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAACCGGTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGGAACACCA
TCAAGAGAACCGGTACGACCTGGCCCGCTGCGCTGATAGGTCAACAGAAAGCTCGGGCTGCATGTGC
X 10 20 30 40 50 60 70

ATCATCCGATCACCACGACGCTGACCGGAGAAAGAGCTGCGCCATCGATGATGTGCTGCGCTGC
GCGACCCGATCAGCTGCTGCGACGCTGACGAAAGAGCTGCTGCGCCATCGAATATCTGCTGCTTGC
X 80 90 100 110 120 130 140

ACGAGGCCAGGACGATGACGCTGCGCGGCGGAGTGAAGTGAAGAAACCGACACAT
ACGAGGCCAGGACGATGACGCTGCGCGGCGGAGTGAAGTGAAGAAACCGACACAT
X 150 160 170 180 190 200

12. US-09-697-123b-8 (1-208)
US-09-697-123b-2 Sequence 22, Application US/09697123b

Initial Score = 181 Optimized Score = 181 Significance = 0.55
Residue Identity = 87% Matches = 181 Mismatches = 27
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAACCGGTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGGAACACCA
TCAAGAGAACCGGTACGACCTGGCCCGCTGCGCTGATAGGTCAACAGAAAGCTCGGGCTGCATGTGC
X 10 20 30 40 50 60 70

ATCATCCGATCACCACGACGCTGACCGGAGAAAGAGCTGCGCCATCGATGATGTGCTGCGCTGC
GCGACCCGATCAGCTGCTGCGACGCTGACGAAAGAGCTGCTGCGCCATCGAATATCTGCTGCTTGC
X 80 90 100 110 120 130 140

ACGAGGCCAGGACGATGACGCTGCGCGGCGGAGTGAAGTGAAGAAACCGACACAT
ACGAGGCCAGGACGATGATGTGCTGCGCGGCGGAGTGAAGTGAAGAAACCGACACAT
X 150 160 170 180 190 200

13. US-09-697-123b-8 (1-208)
US-09-697-123b-2 Sequence 24, Application US/09697123b

Initial Score = 179 Optimized Score = 179 Significance = 0.51
Residue Identity = 86% Matches = 179 Mismatches = 29
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAACCGGTACGACCTGGCCCGCTGCGCCGATCAAGATCAAGAAAGAGCTGGAACACCA
TCAAGAGAACCGGTACGACCTGGCCCGCTGCGCTGATAGGTCAACAGAAAGCTCGGGCTGCATGTGC
X 10 20 30 40 50 60 70

ATCATCCGATCACCACGACGCTGACCGGAGAAAGAGCTGCGCCATCGATGATGTGCTGCGCTGC
AGATGCGCCAAACACGACGCTGACGAAAGAGCTGCTGCGCCATCGAATATCTGCTGCTTGC
X 80 90 100 110 120 130 140

80 90 100 110 120 130 140

19. US-09-697-123B-8 (1-208)
US-09-697-123B-4 Sequence 4, Application US/09697123B

2

Initial Score = 152 Optimized Score = 179 Significance = -0.04
Residue Identity = 88% Matches = 184 Mismatches = 23
Gaps = 1 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCACAGAGAGGCGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
TCACAGAGAGGCGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 10 20 30 40 50 60 70

ATCATTCATACACGACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
CGGATCCCATACACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 10 20 30 40 50 60 70

150 160 170 180 190 200
ACGACGCGGCGGCGACGATGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
ACGACGCGGCGGCGACGATGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 150 160 170 180 190 200

20. US-09-697-123b-8 (1-208)
US-09-697-123b-9 Sequence 9, Application US/09697123b

Initial Score = 138 Optimized Score = 165 Significance = -0.33
Residue Identity = 78% Matches = 175 Mismatches = 33
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCACAGAGAGGCGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
TCACAGAGAGGCGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 10 20 30 40 50 60 70

ATCATTCATACACGACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
CGGATCCCATACACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 10 20 30 40 50 60 70

150 160 170 180 190 200
ACGACGCGGCGGCGACGATGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
ACGACGCGGCGGCGACGATGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 150 160 170 180 190 200

ACGACAT
ACGACAT
ACGACAT
X 220

21. US-09-697-123b-8 (1-208)
US-09-697-123b-7 Sequence 7, Application US/09697123b

Initial Score = 127 Optimized Score = 164 Significance = -0.55
Residue Identity = 79% Matches = 171 Mismatches = 37
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCACAGAGAGGCGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
TCACAGAGAGGCGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 10 20 30 40 50 60 70

ACACCAATCATCCGATACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
CCACCGGCGCTGCTGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 80 90 100 110 120 130 140

140 150 160 170 180 190 200
GCCTCAGAGAGGCGGCGACGATGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
GCCTCAGAGAGGCGGCGACGATGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 140 150 160 170 180 190 200

22. US-09-697-123b-8 (1-208)
US-09-697-123b-2 Sequence 21, Application US/09697123b

Initial Score = 122 Optimized Score = 170 Significance = -0.66
Residue Identity = 82% Matches = 177 Mismatches = 31
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCACAGAGAGGCGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
TCACAGAGAGGCGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 10 20 30 40 50 60 70

ATCATTC-----GATCACCACGACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
CCAATCCGCGCTCAGGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 80 90 100 110 120 130 140

140 150 160 170 180 190 200
GCCTCAGAGAGGCGGCGACGATGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
GCCTCAGAGAGGCGGCGACGATGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 140 150 160 170 180 190 200

23. US-09-697-123b-8 (1-208)
US-09-697-123b-1 Sequence 14, Application US/09697123b

Initial Score = 118 Optimized Score = 168 Significance = -0.74
Residue Identity = 81% Matches = 175 Mismatches = 33
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCACAGAGAGGCGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
TCACAGAGAGGCGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 10 20 30 40 50 60 70

ATCATTCG-----ATCACCACGACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
CCGATCCGCGCTCAGGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 80 90 100 110 120 130 140

140 150 160 170 180 190 200
GCCTCAGAGAGGCGGCGACGATGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
GCCTCAGAGAGGCGGCGACGATGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 140 150 160 170 180 190 200

24. US-09-697-123b-8 (1-208)
US-09-697-123b-1 Sequence 18, Application US/09697123b

Initial Score = 117 Optimized Score = 164 Significance = -0.76
Residue Identity = 80% Matches = 170 Mismatches = 38
Gaps = 3 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCACAGAGAGGCGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
TCACAGAGAGGCGTACGACGCTGGCCCGCGCGATACAGGTCAACAAGAGAGGCGCTGAACACCA
X 10 20 30 40 50 60 70

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Query sequence being compared:	US-09-657-123B-7 (1-214)
Number of sequences searched:	26
Number of scores above cutoff:	25

Results of the initial comparison of US-09-697-123B-7 (1-214) with:
File : US09697123B.seq

100-
-
N
-
U 50-
M
B
E
R
O
F
S
E
O
U
E
N
C
E
S
0-
SCORE
0
STDEV
-2
24
48
-1
71
95
0
119
143
1
166
2
190
214

PARAMETERS	
Unary	4
Similarity matrix	30
Mismatch penalty	30
Gap penalty	208
Gap size penalty	
Cutoff score	
Randomization group	

Scores:	Mean	Median	Standard Deviation
---------	------	--------	--------------------

	CPU	Total Elapsed
Times:		

Number of residues: 50

Number of scores above cutoff:

The scores below are sorted by initial score. Significance is calculated based on initial score.

Sequence Name	Description	Length	Init. Opt. Score	Sig. Frame
---------------	-------------	--------	------------------	------------

Sequence Name	Description	Length	Int. Score	Opt. Score	Sig. Frame
---------------	-------------	--------	------------	------------	------------

	*** 1 standard deviation above mean ***	197	1.90	0
2. US-09-697-123B-2 Sequence 21, Application	214	197	1.90	0
3. US-09-697-123B-1 Sequence 14, Application	214	172	1.31	0
4. US-09-697-123B-1 Sequence 17, Application	208	137	0.47	0
5. US-09-697-123B-1 Sequence 15, Application	208	131	0.33	0
6. US-09-697-123B-2 Sequence 2, Application U	208	129	0.28	0
7. US-09-697-123B-1 Sequence 1, Application U	208	129	0.28	0
8. US-09-697-123B-8 Sequence 8, Application U	208	127	0.24	0
9. US-09-697-123B-3 Sequence 3, Application U	208	125	0.19	0
10. US-09-697-123B-2 Sequence 24, Application	208	123	0.14	0
11. US-09-697-123B-1 Sequence 18, Application	211	122	0.12	0
12. US-09-697-123B-1 Sequence 16, Application	208	121	0.09	0
13. US-09-697-123B-6 Sequence 6, Application U	208	121	0.09	0
14. US-09-697-123B-5 Sequence 5, Application U	208	121	0.09	0
15. US-09-697-123B-2 Sequence 22, Application	208	120	0.07	0
16. US-09-697-123B-2 Sequence 23, Application	208	117	0.00	0
17. US-09-697-123B-1 Sequence 19, Application	208	115	-0.05	0
18. US-09-697-123B-2 Sequence 20, Application	205	109	-0.19	0
19. US-09-697-123B-1 Sequence 11, Application	208	108	-0.21	0
20. US-09-697-123B-1 Sequence 10, Application	208	107	-0.24	0
21. US-09-697-123B-1 Sequence 12, Application	207	101	-0.38	0
22. US-09-697-123B-4 Sequence 4, Application U	207	92	-0.59	0
23. US-09-697-123B-1 Sequence 13, Application	223	92	-0.59	0
24. US-09-697-123B-9 Sequence 9, Application U	223	91	-0.62	0
25. US-09-697-123B-2 Sequence 25, Application	19	19	-2.33	0

1. US-09-697-123B-7 (1-214)
US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score	-	214	Optimized Score	-	214	Significance	-	2.30
Residue Identity	-	100%	Matches	-	214	Mismatches	-	0
Gaps	-	0	Conservative Substitutions	-			-	0

X	10	20	30	40	50	60	70
TCAGGAGAAAGCGCTACGACCTGGCCCGGGTGGCCGGTACAAGGTGAAACAAGACCTGGCTTGGCCGG							
TCAGGAGAAAGCGCTACGACCTGGCCCGGGTGGCCGGTACAAGGTGAAACAAGACCTGGCTTGGCCGG							
TCAAGGAGAAAGCGCTACGACCTGGCCCGGGTGGCCGGTACAAGGTGAAACAAGACCTGGCTTGGCCGG							

CCAAACCCGGCTCTGGTATCTGACACCAACGCTCAACCGAGGAACGTCGTGCCACCATGGGTACTGTGTG
80 90 100 110 120 130 140
CCAAACCCGGCTCTGGTATCTGACACCAACGCTCAACCGAGGAACGTCGTGCCACCATGGGTACTGTGTG
80 90 100 110 120 130 140

150	160	170	180	190	200	210	X
GCTTGCACGAGGGCGACACCATATACCGCCCGGGGGGCTCGAGGTCCCGGTTCGAGTCCACACACT							
150	160	170	180	190	200	210	X
GCTTGCACGAGGGCGACACCATATACCGCCCGGGGGGCTCGAGGTCCCGGTTCGAGTCCACACACT							

2. US-09-697-123B-7 (1-214)
US-09-697-123B-2 Sequence 21, Application US/09697123B

Initial Score	-	197	Optimized Score	-	197	Significance	-	1.90
Residue Identity	-	92%	Matches	-	197	Mismatches	-	17
Gaps	-	0	Conservative Substitutions	-			-	0

X 10 20 30 40 50 60 70
 TCAAGGAAAGCCCTACGACCTTGGCCCGCTGGGCGGTACAAAGTGAACAAGAACTTGGCTTGGCGGTG
 TCAAGGAAAGCCGTACGATCTTGGCCCGCTGGGTGCGTACAAAGTGAACAAGAAAGCTTGGCTTGGCGGTG

[illegible]

3. US-09-697-123B-7 (1-214)
US-09-697-123B-1 Sequence 14, Application US/09697123B

[illegible]

CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 80 90 100 110 120 130 140
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 150 160 170 180 190 200 210
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 220 230 240 250 260 270 280
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 290 300 310 320 330 340 350
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 360 370 380 390 400 410 420
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 430 440 450 460 470 480 490
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 500 510 520 530 540 550 560
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 570 580 590 600 610 620 630
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 640 650 660 670 680 690 700
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 710 720 730 740 750 760 770
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 780 790 800 810 820 830 840
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 850 860 870 880 890 900 910
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 920 930 940 950 960 970 980
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 990 1000 1010 1020 1030 1040 1050
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1060 1070 1080 1090 1100 1110 1120
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1130 1140 1150 1160 1170 1180 1190
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1200 1210 1220 1230 1240 1250 1260
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1270 1280 1290 1300 1310 1320 1330
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1340 1350 1360 1370 1380 1390 1400
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1410 1420 1430 1440 1450 1460 1470
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1480 1490 1500 1510 1520 1530 1540
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1550 1560 1570 1580 1590 1600 1610
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1620 1630 1640 1650 1660 1670 1680
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1690 1700 1710 1720 1730 1740 1750
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1760 1770 1780 1790 1800 1810 1820
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1830 1840 1850 1860 1870 1880 1890
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1900 1910 1920 1930 1940 1950 1960
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 1970 1980 1990 2000 2010 2020 2030
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2040 2050 2060 2070 2080 2090 2100
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2110 2120 2130 2140 2150 2160 2170
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2180 2190 2200 2210 2220 2230 2240
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2250 2260 2270 2280 2290 2300 2310
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2320 2330 2340 2350 2360 2370 2380
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2390 2400 2410 2420 2430 2440 2450
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2460 2470 2480 2490 2500 2510 2520
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2530 2540 2550 2560 2570 2580 2590
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2600 2610 2620 2630 2640 2650 2660
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2670 2680 2690 2700 2710 2720 2730
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2740 2750 2760 2770 2780 2790 2800
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2810 2820 2830 2840 2850 2860 2870
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2880 2890 2900 2910 2920 2930 2940
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 2950 2960 2970 2980 2990 3000 3010
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 3020 3030 3040 3050 3060 3070 3080
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 3090 3100 3110 3120 3130 3140 3150
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 3160 3170 3180 3190 3200 3210 3220
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 3230 3240 3250 3260 3270 3280 3290
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 3300 3310 3320 3330 3340 3350 3360
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 3370 3380 3390 3400 3410 3420 3430
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 3440 3450 3460 3470 3480 3490 3500
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 3510 3520 3530 3540 3550 3560 3570
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 3580 3590 3600 3610 3620 3630 3640
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 3650 3660 3670 3680 3690 3700 3710
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 3720 3730 3740 3750 3760 3770 3780
 CCAACCCGGCTGTGGTGCATCCACACGCTCACCCAGGAGACGTGTGCCACCATGGGGTACCTGTGCT
 3790 3800 38

150	160	170	180	190	200	210	X
GCCTGCACGAGGGCCGAGACCCAGATGACCGCCCCGGCGCTTGAGATGCCGTTGAGTGCAGACAT							
150	160	170	180	190	200	210	X
GCCTGCACGAGGGCCGAGACCCAGATGACCGCCCCGGCGCTTGAGATGCCGTTGAGTGCAGACAT							
150	160	170	180	190	200	210	X
GCCTGCACGAGGGCCGAGACCCAGATGACCGCCCCGGCGCTTGAGATGCCGTTGAGTGCAGACAT							

4. US-09-697-123B-7 (1-214)
US-09-697-123B-1 Sequence 17, Application US/09697123B

Initial Score =	137	Optimized Score =	165	Significance =	0.47
Residue Identity =	80%	Matches	= 172	Mismatches	= 36
Gaps	= 6	Conservative Substitutions			= 0

TCAGAGAGAAAGCGCTACGACCTGGCCCCGCTGGCCCGGTACAAAGGTGAACAAAGCTGGGGTTCGGCGGTTG
 TCAGAGAGAAAGCGCTACGACCTGGCGCGGGTGGCCCGCTACAAAGTTCAACAGAACCTGGCCT-----GAACTG

CCAAACCCGCTGTGTGCTGCCACACGCTCACCGAGAGACGTCGTGGCCACCATCGGTACTCTGT
80 90 100 110 120 130 140
ACACCCGTCGCCGTCACAGCAGCACCTGTGACCGAAGAGAGCGTCGGCCACCATCGAGTACTCTGT
70 80 90 100 110 120 130

[illegible]

5. US-09-697-123B-7 (1-214)
US-09-697-123B-1 Sequence 15, Application US/09697123B

Initial Score	=	131	Optimized Score	=	170	Significance	=	0.33
Residue Identity	=	82%	Matches	=	177	Mismatches	=	31
Gaps	=	6	Conservative Substitutions				=	0

X 10 20 30 40 50 60 70
 TCAAGGAGAGCGCTACGACCGCGCCGCTGGCGGCTACAGTGAACAAAGCTGGCTCTTGGCGGTG
 X 10 20 30 40 50 60 70
 TCAAGGAGAGCGCTACGACCTGGCGGTGGCGCTACAGTGAACAAAGCTGGCTCTTGGCGGTG

CCACCGGCTGTGGTACTCCACAGCTGCACGAGAGACGTGCCACCATTGGTTACTCTTC
80 90 100 110 120 130 140
| | | | |
GTGAGCC-----GATCACACAGCTGCACCGTGACCAGSAGACCTGTCTGCCCATCAGTAATCACTGTGC
80 90 100 110 120 130

150 160 170 180 190 200 210 X
 GCCGCGACGAGGGGCCAGACCAGATGACCCCGCGCGGCTTCGAGTCCGCTGAGGTGCAGACAT
 |||||
 GCGTGCACGAGGCTCAGCCACAGATGACCTCCCGGCGGAGTCTGAGTCCCGCTGAGGTGCAGACAT
 |||||
 150 160 170 180 190 200 X
 |||||

6. US-09-697-123B-7 (1-214)
US-09-697-123B-2 Sequence 2, Application US/09697123B

Initial Score	=	129	Optimized Score	=	166	Significance	=	0.28
Residue Identity	=	80%	Matches	=	173	Mismatches	=	35
Gaps	=	6	Conservative Substitutions				=	0

X	10	20	30	40	50	60	70
TCAGAGAGAGCGCTACGACCTGCGCCGGTGGCCGGGTACAAAGTGAACAAAGATCGGCTTTGGCGGTG							
TCAGAGAGAGCGCTACGACCTGCGCCGGTGGCCGGGTACAAAGTGAACAAAGATCGGCTTTGAACGTGCG							
X	10	20	30	40	50	60	70

[illegible]

150	150	170	180	200	210	X
GCCCTGACAGGAGGCGACACACGATGACACCGCCGCGCGCTGAGAGTCCGGGTGAGGTGAGTGCAGACAT						
150	150	170	180	200	210	X
GGCTTGACAGGAGGTACAGTGGCGATGACCGTTCGCCGCGCGCGAGGTGCGGTGAGACCGACGACAT						
150	150	170	180	200	210	X
GGCTTGACAGGAGGTACAGTGGCGATGACCGTTCGCCGCGCGCGAGGTGCGGTGAGACCGACGACAT						

7. US-09-697-123B-7 (1-214)
US-09-697-123B-1 Sequence 1 And:oct16 us 0000071335

Initial Score	=	129	Optimized Score	=	168	Significance	=	0.28
Residue Identity	=	81%	Matches	=	175	Mismatches	=	33
Gaps	=	6	Conservative Substitutions	=			=	0

[illegible]

CCAACCCGGCTTGTGTACTGCGCCACGCTTACCGAGAGAAAGTGTGCGCACCATCGGGTACCTGGTC
80 90 100 110 120 130 140
GG-----GATCCCATCAGCAGGTCCAGCGTACGAGAGAAAGTGTGCGCACCATCGAGTACCTGGTCC
80 90 100 110 120 130

120 130 140 150 160 170 180 190 200 210
 GCGCTTACGAGGAGGACAGACACCATGTGACCGCCCGGAGGCTCGAGTCCGCGTCAAGCAGCAT
 GCGCTCAGAGAGGAGCAGACACCATGTGACCGCTCCCGGAGGCTCGAGTCCGCGTCAAGCAGCAT
 GCGCTCAGAGAGGAGCAGACACCATGTGACCGCTCCCGGAGGCTCGAGTCCGCGTCAAGCAGCAT
 150 160 170 180 190 200 X
 X

8. US-09-697-123B-7 (1-214)

CCAAACCGGCTGTGGTACATGCCACACACCTTACCAGAGAAGACGTGTGTCCCAACCATTCGGGTACTGTGTGC
80 90 100 110 120 130 140
| | | | | | | | | | | | | | |
GCAGGCC-----CATTACGTGTGTGACCTGACCGAAGAAGACGTGTGTGGCCACCATTCATAATATGTGGTCC

14. US-09-697-123B-7 (1-214)
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 121 Optimized Score = 159 Significance = 0.09
Residue Identity = 77% Matches = 166 Mismatches = 42
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAAGCCCTACGACCTGCGCCCGCGTGGCCGCTACAGGTGAAGAAGCTGCTTGGCGGTG
|||||
TCAAGGAGAAGCCCTACGACCTGCGCCCGCGTGGCCGCTACAGGTGAAGAAGCTGCTTGGCGGTG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CCAACCCGGCTGTGGTACCTGCACACCTCACCAGGAGAGACGTCTGCCACACATCGGCTACCTGTGC
|||||
CCAACCCGGCTGTGGTACCTGCACACCTCACCAGGAGAGACGTCTGCCACACATCGGCTACCTGTGC
GCCAGCC-----CATCAGCTGTGACGCTGACCGAAGAACGTCGTGGCCACATCGAATATGTGTGCC
80 90 100 110 120 130

150 160 170 180 190 200 210
GCTTCGACGAGGCGCCGACCCAGATGACCGCCCGCGGCGCTGAGGTCCCGGTCCAGGTTCAGACAT
|||||
GCTTCGACGAGGCGCCGACCCAGATGACCGCCCGCGGCGCTGAGGTCCCGGTCCAGGTTCAGACAT
X 150 160 170 180 190 200 210

15. US-09-697-123B-7 (1-214)
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 120 Optimized Score = 158 Significance = 0.07
Residue Identity = 77% Matches = 165 Mismatches = 43
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAAGCCCTACGACCTGCGCCCGCGTGGCCGCTACAGGTGAAGAAGCTGCTTGGCGGTG
|||||
TCAAGGAGAAGCCCTACGACCTGCGCCCGCGTGGCCGCTACAGGTGAAGAAGCTGCTTGGCGGTG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CCAACCCGGCTGTGGTACCTGCACACCTCACCAGGAGAGACGTCTGCCACACATCGGCTACCTGTGC
|||||
CCAACCCGGCTGTGGTACCTGCACACCTCACCAGGAGAGACGTCTGCCACACATCGGCTACCTGTGC
GCCAGCC-----CATCAGCTGTGACGCTGACCGAAGAACGTCGTGGCCACATCGAATATGTGTGCC
80 90 100 110 120 130

150 160 170 180 190 200 210
GCTTCGACGAGGCGCCAGACGATGACCGCCCGCGGCGCTGAGGTCCCGGTCCAGGTTCAGACAT
|||||
GCTTCGACGAGGCGCCAGACGATGACCGCCCGCGGCGCTGAGGTCCCGGTCCAGGTTCAGACAT
X 150 160 170 180 190 200 210

16. US-09-697-123B-7 (1-214)
US-09-697-123B-2 Sequence 23, Application US/09697123B

Initial Score = 117 Optimized Score = 155 Significance = 0.00
Residue Identity = 75% Matches = 162 Mismatches = 46
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAAGCCCTACGACCTGCGCCCGCGTGGCCGCTACAGGTGAAGAAGCTGCTTGGCGGTG
|||||
TCAAGGAGAAGCCCTACGACCTGCGCCCGCGTGGCCGCTACAGGTGAAGAAGCTGCTTGGCGGTG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CCAACCCGGCTGTGGTACCTGCACACCTCACCAGGAGAGACGTCTGCCACACATCGGCTACCTGTGC
|||||
CCAACCCGGCTGTGGTACCTGCACACCTCACCAGGAGAGACGTCTGCCACACATCGGCTACCTGTGC
GTGAGCC-----GATCAGACCTGACCGCTACCCAGAGAGACGTCTGCCACACATCGGCTACCTGTGC
80 90 100 110 120 130

150 160 170 180 190 200 210
GCTTCGACGAGGCGCCAGACGATGACCGCCCGCGGCGCTGAGGTCCCGGTCCAGGTTCAGACAT
|||||
GCTTCGACGAGGCGCCAGACGATGACCGCCCGCGGCGCTGAGGTCCCGGTCCAGGTTCAGACAT
X 150 160 170 180 190 200 210

17. US-09-697-123B-7 (1-214)
US-09-697-123B-1 Sequence 19, Application US/09697123B

Initial Score = 115 Optimized Score = 177 Significance = -0.05
Residue Identity = 85% Matches = 184 Mismatches = 24
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAAGCCCTACGACCTGCGCCCGCGTGGCCGCTACAGGTGAAGAAGCTGCTTGGCGGTG
|||||
TCAAGGAGAAGCCCTACGACCTGCGCCCGCGTGGCCGCTACAGGTGAAGAAGCTGCTTGGCGGTG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CCAACCCGGCTGTGGTACCTGCACACCTCACCAGGAGAGACGTCTGCCACACATCGGCTACCTGTGC
|||||
CCAACCCGGCTGTGGTACCTGCACACCTCACCAGGAGAGACGTCTGCCACACATCGGCTACCTGTGC
GCCAGCC-----GATCAGCTGTGACGCTGACCGAAGAACGTCGTGGCCACATCGAATATGTGTGCC
80 90 100 110 120 130

150 160 170 180 190 200 210
GCTTCGACGAGGCGCCAGACGATGACCGCCCGCGGCGCTGAGGTCCCGGTCCAGGTTCAGACAT
|||||
GCTTCGACGAGGCGCCAGACGATGACCGCCCGCGGCGCTGAGGTCCCGGTCCAGGTTCAGACAT
X 150 160 170 180 190 200 210

18. US-09-697-123B-7 (1-214)
US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score = 109 Optimized Score = 158 Significance = -0.19
Residue Identity = 79% Matches = 170 Mismatches = 35
Gaps = 9 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAAGCCCTACGACCTGCGCCCGCGTGGCCGCTACAGGTGAAGAAGCTGCTTGGCGGTG
|||||
TCAAGGAGAAGCCCTACGACCTGCGCCCGCGTGGCCGCTACAGGTGAAGAAGCTGCTTGGCGGTG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CCAACCCGGCTGTGGTACCTGCACACCTCACCAGGAGAGACGTCTGCCACACATCGGCTACCTGTGC
|||||
CCAACCCGGCTGTGGTACCTGCACACCTCACCAGGAGAGACGTCTGCCACACATCGGCTACCTGTGC
GCCAGCC-----GATCAGACCTGACCGCTACCCAGAGAGACGTCTGCCACACATCGGCTACCTGTGC
80 90 100 110 120 130

150 160 170 180 190 200 210
GCTTCGACGAGGCGCCAGACGATGACCGCCCGCGGCGCTGAGGTCCCGGTCCAGGTTCAGACAT
|||||
GCTTCGACGAGGCGCCAGACGATGACCGCCCGCGGCGCTGAGGTCCCGGTCCAGGTTCAGACAT
X 150 160 170 180 190 200 210

19. US-09-697-123B-7 (1-214)
US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 108 Optimized Score = 158 Significance = -0.21
Residue Identity = 81% Matches = 175 Mismatches = 33
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAACAAAGAGCTGGGCTTGGCGGTG
|||||
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAACAAAGAGCTGGGCTTGGCGGTG
X 10 20 30 40 50 60 70
CCAAACCGGCTCTGTGATGCTGCCACCGCTACCGAGAGAGCTGCTGCGACATCTGGGTACCTGTGCG
|||||
GCCAGCC-----CATCACACAGCTGACGCTGACCGAGAGAGCTGCTGCGACATCTGAAATACCTGTGCG
80 90 100 110 120 130 140
GCGTCCACGAGGCGCCAGACAGATGACCGCCCGCGGTGACAGTCCCGGTGAGTCCGCTGAGTCCAGCAT
|||||
GCTTGACAGAGGCGCCAGACAGGATGACCGCTGCGGGGGGTGTGAGGTGCGGCTGAGAGCAGACAT
140 150 160 170 180 190 200 X

20. US-09-697-123B-7 (1-214)

US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 107 Optimized Score = 167 Significance = -0.24
Residue Identity = 81% Matches = 174 Mismatches = 34
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAACAAAGCTGCTTGGCGGTG
|||||
TCAAGGAGAGGCGTACGACCTGCGCTCCGCTGGGTGCGTACAAAGTGAACAAAGCTGCTTGGCGGTG
X 10 20 30 40 50 60 70
CCAAACCGGCTCTGTGATGCTGCCACCGCTACCGAGAGAGCTGCTGCGACATCTGGGTACCTGTGCG
|||||
GCCAGCC-----CATCACACAGCTGACGCTGACCGAGAGAGCTGCTGCGACATCTGAAATACCTGTGCG
80 90 100 110 120 130 140
GCGTCCACGAGGCGCCAGACAGATGACCGCCCGCGGTGACAGTCCCGGTGAGTCCGCTGAGTCCAGCAT
|||||
GCTTGACAGAGGCGCCAGACAGGATGACCGCTGCGGGGGGTGTGAGGTGCGGCTGAGAGCAGACAT
140 150 160 170 180 190 200 X

21. US-09-697-123B-7 (1-214)

US-09-697-123B-1 Sequence 12, Application US/09697123B

Initial Score = 101 Optimized Score = 158 Significance = -0.38
Residue Identity = 79% Matches = 170 Mismatches = 37
Gaps = 7 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAACAAAGCTGCTTGGCGGTG
|||||
TCAAGGAGAGGCGTACGACCTGCGCTCCGCTGGGTGCGTACAAAGTGAACAAAGCTGCTTGGCGGTG
X 10 20 30 40 50 60 70
CCAAACCGGCTCTGTGATGCTGCCACCGCTACCGAGAGAGCTGCTGCGACATCTGGGTACCTGTGCG
|||||
GCCAGCC-----CATCACACAGCTGACGCTGACCGAGAGAGCTGCTGCGACATCTGAAATACCTGTGCG
80 90 100 110 120 130 140
GCGTCCACGAGGCGCCAGACAGATGACCGCCCGCGGTGACAGTCCCGGTGAGTCCGCTGAGTCCAGCAT
|||||
GCTTGACAGAGGCGCCAGACAGGATGACCGCTGCGGGGGGTGTGAGGTGCGGCTGAGAGCAGACAT
140 150 160 170 180 190 200 X

22. US-09-697-123B-7 (1-214)

US-09-697-123B-4 Sequence 4, Application US/09697123B

Initial Score = 92 Optimized Score = 155 Significance = -0.59
Residue Identity = 78% Matches = 167 Mismatches = 40
Gaps = 7 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAACAAAGCTGCTTGGCGGTG
|||||
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAACAAAGCTGCTTGGCGGTG
X 10 20 30 40 50 60 70
CCAAACCGGCTCTGTGATGCTGCCACCGCTACCGAGAGAGCTGCTGCGACATCTGGGTACCTGTGCG
|||||
GCCAGCC-----CATCACACAGCTGACGCTGACCGAGAGAGCTGCTGCGACATCTGAAATACCTGTGCG
80 90 100 110 120 130 140
GCGTCCACGAGGCGCCAGACAGATGACCGCCCGCGGTGACAGTCCCGGTGAGTCCGCTGAGTCCAGCAT
|||||
GCTTGACAGAGGCGCCAGACAGGATGACCGCTGCGGGGGGTGTGAGGTGCGGCTGAGAGCAGACAT
140 150 160 170 180 190 200 X

23. US-09-697-123B-7 (1-214)

US-09-697-123B-1 Sequence 13, Application US/09697123B

Initial Score = 92 Optimized Score = 150 Significance = -0.59
Residue Identity = 74% Matches = 171 Mismatches = 37
Gaps = 21 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAACAAAGCTGCTTGGCGGTG
|||||
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAACAAAGCTGCTTGGCGGTG
X 10 20 30 40 50 60 70
CCAAACCGGCTCTGTGATGCTGCCACCGCTACCGAGAGAGCTGCTGCGACATCTGGGTACCTGTGCG
|||||
ACACCGATCATCTGATACACACACACACCTGACCGAGAGAGCTGCTGCGACATCTGAAATACCTGTGCG
70 80 90 100 110 120 130
GCGTCCACGAGGCGCCAGACAGATGACCGCCCGCGGTGACAGTCCCGGTGAGTCCGCTGAGTCCAGCAT
140 150 160 170 180 190 200 210
AGGTGACGACAT
|||||
AAACGACGACAT
220 X

24. US-09-697-123B-7 (1-214)

US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 91 Optimized Score = 154 Significance = -0.62
Residue Identity = 74% Matches = 170 Mismatches = 38
Gaps = 21 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAACAAAGCTGCTTGGCGGTG
|||||
TCAAGGAGAGGCGTACGACCTGCGCCCGCGGTGACAAAGTGAACAAAGCTGCTTGGCGGTG
X 10 20 30 40 50 60 70
CCAAACCGGCTCTGTGATGCTGCCACCGCTACCGAGAGAGCTGCTGCGACATCTGGGTACCTGTGCG
|||||
GCCAGCC-----CATCACAGCTGCTGACCGAGAGAGCTGCTGCGACATCTGAAATACCTGTGCG
80 90 100 110 120 130 140
GCGTCCACGAGGCGCCAGACAGATGACCGCCCGCGGTGACAGTCCCGGTGAGTCCGCTGAGTCCAGCAT
|||||

> 0 <
0100 Intelligenetics
> 0 <
FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-09-697-123b-6.res made by shanley on Wed 13 Nov 102 14:04:08-PRF.

Query sequence being compared: US-09-697-123b-6 (1-208)
Number of sequences searched: 26
Number of scores above cutoff: 25

Results of the initial comparison of US-09-697-123b-6 (1-208) with:
File: US09697123b.seq

```

100-
N -
U -
B -
M -
E -
R -
O -
F 10-
S -
E 5-
U -
U -
E -
C -
S 0-
SCORE 0 231 46 69 92 116 139 162 185 208
STDEV -2 -1 0 1

```

PARAMETERS

Similarity matrix Unitary K-tuple
Mismatch penalty 1 Joining penalty 4
Gap penalty 5.00 Window size 30
Gap size penalty 0.33
Cutoff score 12
Randomization group 0

SEARCH STATISTICS

Scores: Mean 155 Median 179 Standard Deviation 51.93
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00
Number of residues: 5077
Number of sequences searched: 26
Number of scores above cutoff: 25

The scores below are sorted by initial score.
Significance is calculated based on initial score.

3 100% identical sequences to the query sequence were found:

Sequence Name	Description	Length	Init. Opt. Score	Sig. Frame
---------------	-------------	--------	------------------	------------

The list of other best scores is:

Sequence Name	Description	Length	Init. Opt. Score	Sig. Frame
1. US-09-697-123b-1	Sequence 16, Application	208	208	1.02 0
2. US-09-697-123b-6	Sequence 6, Application U	208	208	1.02 0
3. US-09-697-123b-5	Sequence 5, Application U	208	208	1.02 0
The list of other best scores is:				
**** 1 standard deviation above mean ****				
4. US-09-697-123b-2	Sequence 22, Application	208	207	1.00 0
**** 0 standard deviation from mean ****				
5. US-09-697-123b-1	Sequence 10, Application	208	186	0.60 0
6. US-09-697-123b-1	Sequence 11, Application	208	185	0.58 0
7. US-09-697-123b-3	Sequence 3, Application U	208	185	0.58 0
8. US-09-697-123b-1	Sequence 15, Application	208	184	0.56 0
9. US-09-697-123b-2	Sequence 23, Application	208	182	0.52 0
10. US-09-697-123b-1	Sequence 19, Application	208	182	0.52 0
11. US-09-697-123b-8	Sequence 8, Application U	208	182	0.52 0
12. US-09-697-123b-1	Sequence 1, Application U	208	182	0.52 0
13. US-09-697-123b-2	Sequence 2, Application U	208	181	0.50 0
14. US-09-697-123b-1	Sequence 17, Application	208	178	0.44 0
15. US-09-697-123b-2	Sequence 24, Application	208	172	0.33 0
16. US-09-697-123b-2	Sequence 20, Application	205	161	0.12 0
17. US-09-697-123b-1	Sequence 12, Application	207	157	0.04 0
18. US-09-697-123b-4	Sequence 4, Application U	207	147	0.15 0
19. US-09-697-123b-9	Sequence 9, Application U	223	144	0.21 0
20. US-09-697-123b-1	Sequence 13, Application	223	137	0.35 0
21. US-09-697-123b-7	Sequence 7, Application U	214	121	0.59 0
22. US-09-697-123b-1	Sequence 18, Application	211	113	0.81 0
23. US-09-697-123b-1	Sequence 14, Application	214	110	0.87 0
24. US-09-697-123b-2	Sequence 21, Application	214	109	0.89 0
25. US-09-697-123b-2	Sequence 25, Application	19	19	-2.62 0

1. US-09-697-123b-6 (1-208)

US-09-697-123b-1 Sequence 16, Application US/09697123b

Initial Score	Residue Identity	Optimized Score	Matches	Significance	Mismatches
208	100%	208	208	1.02	0
0	0	0	0	0	0

US-09-697-123b-2 Sequence 22, Application US/09697123b

Initial Score	Residue Identity	Optimized Score	Matches	Significance	Mismatches
208	100%	208	208	1.02	0
0	0	0	0	0	0

US-09-697-123b-3 Sequence 3, Application U

Initial Score	Residue Identity	Optimized Score	Matches	Significance	Mismatches
208	100%	208	208	1.02	0
0	0	0	0	0	0

US-09-697-123b-4 Sequence 4, Application U

Initial Score	Residue Identity	Optimized Score	Matches	Significance	Mismatches
208	100%	208	208	1.02	0
0	0	0	0	0	0

2. US-09-697-123b-6 (1-208)
US-09-697-123b-6 Sequence 6, Application US/09697123b

Initial Score	Residue Identity	Optimized Score	Matches	Significance	Mismatches
208	100%	208	208	1.02	0
0	0	0	0	0	0

X 10 20 30 40 50 60 70

CGGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
|||||
GCGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
80 90 100 110 120 130 140
X

ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
|||||
ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
150 160 170 180 190 200
X

3. US-09-697-123B-6 (1-208)

US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 208 Optimized Score = 208 Significance = 1.02
Residue Identity = 100% Matches = 208 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCGACGAGAACCGCTACGACCTGGCCCGCGCTCGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC
|||||
TCGACGAGAACCGCTACGACCTGGCCCGCGCTCGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC
10 20 30 40 50 60 70
X

CGGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
|||||
CGGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
80 90 100 110 120 130 140
X

ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
|||||
ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
150 160 170 180 190 200
X

4. US-09-697-123B-6 (1-208)

US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 207 Optimized Score = 207 Significance = 1.00
Residue Identity = 99% Matches = 207 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCGACGAGAACCGCTACGACCTGGCCCGCGCTCGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC
|||||
TCGACGAGAACCGCTACGACCTGGCCCGCGCTCGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC
10 20 30 40 50 60 70
X

CGGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
|||||
CGGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
80 90 100 110 120 130 140
X

ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
|||||
ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
150 160 170 180 190 200
X

CGGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
|||||
CGGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
80 90 100 110 120 130 140
X

ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
|||||
ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
150 160 170 180 190 200
X

5. US-09-697-123B-6 (1-208)

US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 186 Optimized Score = 186 Significance = 0.60
Residue Identity = 89% Matches = 186 Mismatches = 22
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCGACGAGAACCGCTACGACCTGGCCCGCGCTCGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC
|||||
TCGACGAGAACCGCTACGACCTGGCCCGCGCTCGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC
10 20 30 40 50 60 70
X

GCGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
|||||
GCGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
80 90 100 110 120 130 140
X

ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
|||||
ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
150 160 170 180 190 200
X

6. US-09-697-123B-6 (1-208)

US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.58
Residue Identity = 88% Matches = 185 Mismatches = 23
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCGACGAGAACCGCTACGACCTGGCCCGCGCTCGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC
|||||
TCGACGAGAACCGCTACGACCTGGCCCGCGCTCGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC
10 20 30 40 50 60 70
X

CGGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
|||||
CGGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
80 90 100 110 120 130 140
X

ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
|||||
ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
150 160 170 180 190 200
X

7. US-09-697-123B-6 (1-208)

US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.58
Residue Identity = 88% Matches = 185 Mismatches = 23
Gaps = 0 Conservative Substitutions = 0

TCGACGAGAACCGCTACGACCTGGCCCGCGCTCGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC
|||||
TCGACGAGAACCGCTACGACCTGGCCCGCGCTCGCTATAGGTCAACAGAAGCTCGGCTGCATGTGC
10 20 30 40 50 60 70
X

CGGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
|||||
CGGAGCCCATACGTCGTCGACGCTGACCCGAAGAAGACGTGTGGCCCATCGCAATATCTGGTCCGCTTGC
80 90 100 110 120 130 140
X

ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
|||||
ACGAGGCTCAGACCAACGATGACGCTTCCGGGGCGGCGTCGAGGTGCCCGGTGGAACCCAGACAT
150 160 170 180 190 200
X

8. US-09-697-123B-6 (1-208)

GCGAGCCCATCTACGTGTCGACGCTTACCAGGAAGAACGTCGTGGCCACCATCGAATATCTTGTCGCCGTTGC
|| ||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
GCAGCGCATCTACCACTCGACGCTCACCAGGAAGACTCTTAACCAACCATCGAGTAACTGTGGCGCGTGC

Initial Score	-	144	Optimized Score	-	170	Significance	-	-0.21
Residue Identity	-	80%	Matches	-	180	Mismatches	-	28
Gaps	-	15	Conservative Substitutions	-			-	0

X 10 20 30 40 50 60 70
 TCAAGGAGAGCGCTACGACCTGGCCCGGCGTGGCTCTATTAAGTTCACACAGAGCTCGGGCTGCATGTCG
 |||||
 TCAGAGAGAGCGCTACGATCTGGCCCGGCGTGGCTCGCTACAGGTGACACAGAGAGCTGGGCTTGGGCGGCA

X 10 20 30 40 50 60 70

GGAGCCCATACGCTGTCGACGCTGACCCAGAGAGAGCTGTCGACCATGCAATATCTGTCGCTTGC
|||||
GCGAGCCCATACGCTGTCGACGCTGACCCAGAGAGAGCTGTCGACCATGCAATATCTGTCGCTTGC
80 90 100 110 120 130 140

ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCGCTCGAGGTGCGCGGTGGAACCGAGACAT X
|||||
ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCGCTCGAGGTGCGCGGTGGAACCGAGACAT
150 160 170 180 190 200

3. US-09-697-123B-5 (1-208)

US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 208 Optimized Score = 208 Significance = 1.02
Residue Identity = 100% Matches = 208 Mismatches = 0
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCCCTAGACCTGCGCCCGCGCTGCTATTAAGCTCAACAGAAGCTCGCGCTGATGTCG
|||||
TCAAGAGAAAGCCCTAGACCTGCGCCCGCGCTGCTATTAAGCTCAACAGAAGCTCGCGCTGATGTCG
10 20 30 40 50 60 70

GGGAGCCCATACGCTGTCGACGCTGACCCAGAGAGAGCTGTCGACCATGCAATATCTGTCGCTTGC
|||||
GGGAGCCCATACGCTGTCGACGCTGACCCAGAGAGAGCTGTCGACCATGCAATATCTGTCGCTTGC
80 90 100 110 120 130 140

ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCGCTCGAGGTGCGCGGTGGAACCGAGACAT X
|||||
ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCGCTCGAGGTGCGCGGTGGAACCGAGACAT
150 160 170 180 190 200

4. US-09-697-123B-5 (1-208)

US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 207 Optimized Score = 207 Significance = 1.00
Residue Identity = 99% Matches = 207 Mismatches = 1
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCCCTAGACCTGCGCCCGCGCTGCTATTAAGCTCAACAGAAGCTCGCGCTGATGTCG
|||||
TCAAGAGAAAGCCCTAGACCTGCGCCCGCGCTGCTATTAAGCTCAACAGAAGCTCGCGCTGATGTCG
10 20 30 40 50 60 70

GGGAGCCCATACGCTGTCGACGCTGACCCAGAGAGAGCTGTCGACCATGCAATATCTGTCGCTTGC
|||||
GGGAGCCCATACGCTGTCGACGCTGACCCAGAGAGAGCTGTCGACCATGCAATATCTGTCGCTTGC
80 90 100 110 120 130 140

ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCGCTCGAGGTGCGCGGTGGAACCGAGACAT X
|||||
ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCGCTCGAGGTGCGCGGTGGAACCGAGACAT
150 160 170 180 190 200

5. US-09-697-123B-5 (1-208)

US-09-697-123B-1 Sequence 10, Application US/09697123B

Initial Score = 186 Optimized Score = 186 Significance = 0.60
Residue Identity = 89% Matches = 186 Mismatches = 22
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70

TCAAGAGAAAGCCCTAGACCTGCGCCCGCGCTGCTATTAAGCTCAACAGAAGCTCGCGCTGATGTCG
|||||
TCAAGAGAAAGCCCTAGACCTGCGCCCGCGCTGCTATTAAGCTCAACAGAAGCTCGCGCTGATGTCG
10 20 30 40 50 60 70

GGGAGCCCATACGCTGTCGACGCTGACCCAGAGAGAGCTGTCGACCATGCAATATCTGTCGCTTGC
|||||
GGGAGCCCATACGCTGTCGACGCTGACCCAGAGAGAGCTGTCGACCATGCAATATCTGTCGCTTGC
80 90 100 110 120 130 140

ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCGCTCGAGGTGCGCGGTGGAACCGAGACAT X
|||||
ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCGCTCGAGGTGCGCGGTGGAACCGAGACAT
150 160 170 180 190 200

6. US-09-697-123B-5 (1-208)

US-09-697-123B-1 Sequence 11, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.58
Residue Identity = 88% Matches = 185 Mismatches = 23
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCCCTAGACCTGCGCCCGCGCTGCTATTAAGCTCAACAGAAGCTCGCGCTGATGTCG
|||||
TCAAGAGAAAGCCCTAGACCTGCGCCCGCGCTGCTATTAAGCTCAACAGAAGCTCGCGCTGATGTCG
10 20 30 40 50 60 70

GGGAGCCCATACGCTGTCGACGCTGACCCAGAGAGAGCTGTCGACCATGCAATATCTGTCGCTTGC
|||||
GGGAGCCCATACGCTGTCGACGCTGACCCAGAGAGAGCTGTCGACCATGCAATATCTGTCGCTTGC
80 90 100 110 120 130 140

ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCGCTCGAGGTGCGCGGTGGAACCGAGACAT X
|||||
ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCGCTCGAGGTGCGCGGTGGAACCGAGACAT
150 160 170 180 190 200

7. US-09-697-123B-5 (1-208)

US-09-697-123B-3 Sequence 3, Application US/09697123B

Initial Score = 185 Optimized Score = 185 Significance = 0.58
Residue Identity = 88% Matches = 185 Mismatches = 23
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAAAGCCCTAGACCTGCGCCCGCGCTGCTATTAAGCTCAACAGAAGCTCGCGCTGATGTCG
|||||
TCAAGAGAAAGCCCTAGACCTGCGCCCGCGCTGCTATTAAGCTCAACAGAAGCTCGCGCTGATGTCG
10 20 30 40 50 60 70

GGGAGCCCATACGCTGTCGACGCTGACCCAGAGAGAGCTGTCGACCATGCAATATCTGTCGCTTGC
|||||
GGGAGCCCATACGCTGTCGACGCTGACCCAGAGAGAGCTGTCGACCATGCAATATCTGTCGCTTGC
80 90 100 110 120 130 140

ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCGCTCGAGGTGCGCGGTGGAACCGAGACAT X
|||||
ACGAGGGTCAGACACCATGACGCTTCCGGGGCGGCGCTCGAGGTGCGCGGTGGAACCGAGACAT
150 160 170 180 190 200

8. US-09-697-123B-5 (1-208)

80 90 100 110 120 130 140
GGGAGCCCATCAGTCGTGTGACGCTACCCGAGAGAGACGTCTGTGGCCACCATCGAATATCTGTGCTCCGTCTTC

Initial Score	=	144	Optimized Score	=	170	Significance	=	-0.21
Residue Identity	=	80%	Matches	=	180	Mismatches	=	28
Gaps	=	15	Conservative Substitutions	=	0			

[illegible]

25. US-09-697-123B-5 (1-208)
US-09-697-123B-2 Sequence 25, Application US/09697123B

Initial Score	=	19	Optimized Score	=	19	Significance	=	-2.62
Residue Identity	=	100%	Matches	=	19	Mismatches	=	0
Gaps	=	0	Conservative Substitutions	=	0			
X	10	X						
TCAAGCAGAACCCCTACGA								
TCAAGCAGAACCCCTACGA								
X	10	X						

Sequence Name	Description	Init. Opt.	Length	Score	Score	Sig.	Frame
---------------	-------------	------------	--------	-------	-------	------	-------

[illegible]

80 90 100 110 120 130 140
 GGGATTCGATATACCAAGTTCGACGCTACCCGAAGAGAGCTGTGCACACATGAGTACTGTGCTCCGCTCC
 |||||
 GGGACCATATACCTCTCGACGCTACCCGAAGAGAGCTGTGCACACATGAGTATCTGTGCTCCGCTTC
 80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCTACGACACGATGA-CGTTCCGGGCGCGACAGCTTCGCGTGAGACGACACAT
|||||
ACGAGGCTACGACACGATGA-CGTTCCGGGCGCGACAGCTTCGCGTGAGACGACACAT
150 160 170 180 190 200 X

14. US-09-697-123B-4 (1-207)
US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 147 Optimized Score = 178 Significance = 0.40
Residue Identity = 87% Matches = 183 Mismatches = 24
Gaps = 1 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAAAGCGCTACGACCTGGCCCGCGCGCTACAGGTCAACAGAAAGCTGCGATGTC
|||||
TCAGAGAAAGCGCTACGACCTGGCCCGCGCGCTACAGGTCAACAGAAAGCTGCGATGTC
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CGGATCCGATCACCAGCTCGACGCTGACCGAAGAGAGAGCTGCGCCACCATCGAGTACCTGTCGCTCC
|||||
CGGATCCGATCACCAGCTCGACGCTGACCGAAGAGAGAGCTGCGCCACCATCGAGTACCTGTCGCTCC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCTACGACACGATGA-CGTTCCGGGCGCGACGAGGTTCGCTGAGACGACGACAT
|||||
ACGAGGCTACGACACGATGA-CGTTCCGGGCGCGACGAGGTTCGCTGAGACGACGACAT
150 160 170 180 190 200 X

15. US-09-697-123B-4 (1-207)
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 147 Optimized Score = 178 Significance = 0.40
Residue Identity = 87% Matches = 183 Mismatches = 24
Gaps = 1 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAAAGCGCTACGACCTGGCCCGCGCGCTACAGGTCAACAGAAAGCTGCGATGTC
|||||
TCAGAGAAAGCGCTACGACCTGGCCCGCGCGCTACAGGTCAACAGAAAGCTGCGATGTC
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CGGATCCGATCACCAGCTCGACGCTGACCGAAGAGAGAGCTGCGCCACCATCGAGTACCTGTCGCTCC
|||||
CGGATCCGATCACCAGCTCGACGCTGACCGAAGAGAGAGCTGCGCCACCATCGAGTACCTGTCGCTCC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCTACGACACGATGA-CGTTCCGGGCGCGACCGAGGTTCGCTGAGAGACGACGACAT
|||||
ACGAGGCTACGACACGATGA-CGTTCCGGGCGCGACCGAGGTTCGCTGAGAGAGACGACAT
150 160 170 180 190 200 X

16. US-09-697-123B-4 (1-207)

US-09-697-123B-1 Sequence 17, Application US/09697123B

Initial Score = 144 Optimized Score = 171 Significance = 0.33
Residue Identity = 84% Matches = 176 Mismatches = 31
Gaps = 1 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAAAGCGCTACGACCTGGCCCGCGCGCTACAGGTCAACAGAAAGCTGCGATGTC
|||||
TCAGAGAAAGCGCTACGACCTGGCCCGCGCGCTACAGGTCAACAGAAAGCTGCGATGTC
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GGGATCCGATCACCAGCTCGACGCTGACCGAAGAGAGAGCTGCGCCACCATCGAGTACCTGTCGCTCC
|||||
GGGATCCGATCACCAGCTCGACGCTGACCGAAGAGAGAGCTGCGCCACCATCGAGTACCTGTCGCTCC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCTACGACACGATGA-CGTTCCGGGCGCGACCGAGGTTCGCTGAGAGACGACGACAT
|||||
ACGAGGCTACGACACGATGA-CGTTCCGGGCGCGACCGAGGTTCGCTGAGAGAGACGACGACAT
150 160 170 180 190 200 X

17. US-09-697-123B-4 (1-207)
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score = 138 Optimized Score = 167 Significance = 0.19
Residue Identity = 82% Matches = 172 Mismatches = 35
Gaps = 1 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAAAGCGCTACGACCTGGCCCGCGCGCTACAGGTCAACAGAAAGCTGCGATGTC
|||||
TCAGAGAAAGCGCTACGACCTGGCCCGCGCGCTACAGGTCAACAGAAAGCTGCGATGTC
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CGGATCCGATCACCAGCTCGACGCTGACCGAAGAGAGAGCTGCGCCACCATCGAGTACCTGTCGCTCC
|||||
CGGATCCGATCACCAGCTCGACGCTGACCGAAGAGAGAGCTGCGCCACCATCGAGTACCTGTCGCTCC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCTACGACACGATGA-CGTTCCGGGCGCGACCGAGGTTCGCTGAGAGACGACGACAT
|||||
ACGAGGCTACGACACGATGA-CGTTCCGGGCGCGACCGAGGTTCGCTGAGAGAGACGACGACAT
150 160 170 180 190 200 X

18. US-09-697-123B-4 (1-207)
US-09-697-123B-1 Sequence 13, Application US/09697123B

Initial Score = 131 Optimized Score = 163 Significance = 0.02
Residue Identity = 79% Matches = 178 Mismatches = 29
Gaps = 16 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAAAGCGCTACGACCTGGCCCGCGCGCTACAGGTCAACAGAAAGCTGCGATGTC
|||||
TCAGAGAAAGCGCTACGACCTGGCCCGCGCGCTACAGGTCAACAGAAAGCTGCGATGTC
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CGGATCCGATCACCAGCTCGACGCTGACCGAAGAGAGAGCTGCGCCACCATCGAGTACCTGTCGCTCC
|||||
CGGATCCGATCACCAGCTCGACGCTGACCGAAGAGAGAGCTGCGCCACCATCGAGTACCTGTCGCTCC
80 90 100 110 120 130 140

AC-----GAGGCTACGACACGATGA-CGTTCCGGGCGCGACCGAGGTTCGCTGAGAGACGCG
|||||
ACGAGGCTACGACACGATGA-CGTTCCGGGCGCGACCGAGGTTCGCTGAGAGAGACGCG
150 160 170 180 190 200 210

X
ACGACAT
ACGACAT
220 X

19. US-09-697-123B-4 (1-207)
US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 127 Optimized Score = 163 Significance = -0.07
Residue Identity = 79% Matches = 178 Mismatches = 29
Gaps = 16 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGAAAGTGGCTGATGTCG
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGAAAGTGGCTGATGTCG
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGAAAGTGGCTGATGTCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
GCGATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
GCGATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
X 10 20 30 40 50 60 70
ACGACAT
ACGACAT
ACGACAT
X 220

150 160 170 180 190 200 210
AC-----GAGGCTACACACATGA---CGTCCGCGCGCGGACCGAGTTCGCGTGGAGACCG
ACGACAT
ACGACAT
ACGACAT
X 220

20. US-09-697-123b-4 (1-207)
US-09-697-123b-1 Sequence 12, Application US/09697123B

Initial Score = 122 Optimized Score = 176 Significance = -0.19
Residue Identity = 89% Matches = 186 Mismatches = 20
Gaps = 2 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
GCGATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
GCGATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
X 10 20 30 40 50 60 70

150 160 170 180 190 200
ACGAGGCTACGACACGATGA-CGTTCCGGGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
ACGAGGCTACGACACGATGA-CGTTCCGGGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
ACGAGGCTACGACACGATGA-CGTTCCGGGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
X 150 160 170 180 190 200

21. US-09-697-123b-4 (1-207)
US-09-697-123b-1 Sequence 18, Application US/09697123B

Initial Score = 106 Optimized Score = 154 Significance = -0.56
Residue Identity = 78% Matches = 165 Mismatches = 42
Gaps = 4 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
GCGATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
GCGATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
X 10 20 30 40 50 60 70

150 160 170 180 190 200
ACGAGGCTACGACACGATGA-CGTTCCGGGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
ACGAGGCTACGACACGATGA-CGTTCCGGGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
ACGAGGCTACGACACGATGA-CGTTCCGGGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
X 150 160 170 180 190 200

80 90 100 110 120 130 140
TCCACGAGGCTACGACACGATGA-CGTTCCGGGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
TCCACGAGGCTACGACACGATGA-CGTTCCGGGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
TCCACGAGGCTACGACACGATGA-CGTTCCGGGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
X 150 160 170 180 190 200

22. US-09-697-123b-4 (1-207)
US-09-697-123b-2 Sequence 14, Application US/09697123B

Initial Score = 97 Optimized Score = 161 Significance = -0.77
Residue Identity = 80% Matches = 173 Mismatches = 34
Gaps = 7 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
GCGATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
GCGATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
X 10 20 30 40 50 60 70

150 160 170 180 190 200 210
GCTTCACAGAGGCTACGACACGATGAC-GTTCGCGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
GCTTCACAGAGGCTACGACACGATGAC-GTTCGCGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
GCTTCACAGAGGCTACGACACGATGAC-GTTCGCGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
X 150 160 170 180 190 200 210

23. US-09-697-123b-4 (1-207)
US-09-697-123b-7 Sequence 7, Application US/09697123B

Initial Score = 92 Optimized Score = 155 Significance = -0.89
Residue Identity = 78% Matches = 167 Mismatches = 40
Gaps = 7 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130
GC-----GATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
GC-----GATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
GC-----GATCCGATCACCACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
X 150 160 170 180 190 200

24. US-09-697-123b-4 (1-207)
US-09-697-123b-2 Sequence 21, Application US/09697123B

Initial Score = 82 Optimized Score = 157 Significance = -1.12
Residue Identity = 78% Matches = 169 Mismatches = 38
Gaps = 7 Conservative Substitutions = 0

X
10 20 30 40 50 60 70
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
TCAAGAGAAAGCCCTACGACCTGCGCCGCTGCGCCGCTACAAAGTCAACAAAGTGGCTGATGTCG
X 10 20 30 40 50 60 70

150 160 170 180 190 200 210
GCTTCACAGAGGCTACGACACGATGAC-GTTCGCGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
GCTTCACAGAGGCTACGACACGATGAC-GTTCGCGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
GCTTCACAGAGGCTACGACACGATGAC-GTTCGCGCGGACCGAGGTTCCGGTGGAGACCGACGACAT
X 150 160 170 180 190 200 210

x	10	20	30	40	50	60	70
---	----	----	----	----	----	----	----

8. US-09-697-123B-3 (1-208)
US-09-697-123B-5 Sequence 5, Application US/09697123B

GCGATCCGATCACCACTTCACGCTGACCAGGAAGAAGCCTGCCTGCCACCATCGAATTACTTGGTCGCTCT
80 90 100 110 120 130 140

GCAGCGCATCAGCTGCTGACGTCTGACCGAAGACGCTGCCTGCCACCATCGAATTACTTGGTCGCTCT
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGTCAGACACCGTCCGGGCGGACCGAGGTTCCGGTGGAGCCGACGACAT
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
ACGAGGCGACACCGACGATGACCTCCCGGCGGTGAGGTCCGGTGAAGTGAACGACAT
150 160 170 180 190 200 X

14. US-09-697-123b-3 (1-208)

US-09-697-123b-1 Sequence 17, Application US/09697123b

Initial Score = 179 Optimized Score = 179 Significance = 0.47
Residue Identity = 86% Matches = 179 Mismatches = 29
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAACCCCTACGACCTGGCCCGGTGCGCCCTACAGGTCAACAAGAGCTGGCCCTGACAGTCG
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
TCAAGAGAACCCCTACGACCTGGCCCGGTGCGCCCTACAGGTCAACAAGAGCTGGCCCTGACAGTCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
CGCATCCGATACGACGCTCCACGCTGACCGAAGAGCTGTCGCGCATCGATGACCTGCTCGCTGC
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
CGTCCCGGATACGACGACCTGCTGACCGAAGAGAGAGCTGTCGCCACCATCGATGACCTGCTCGCTGC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGTCAGACACCGTACGACCTGGCCCGGCGGACCGAGGTTCCGGTGGAGCCGACGACAT
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
ACGAGGCGACACCGACGATGACCTCCCGGCGGTGAGGTCCGGTGAAGCCGACGACAT
150 160 170 180 190 200 X

15. US-09-697-123b-3 (1-208)

US-09-697-123b-2 Sequence 24, Application US/09697123b

Initial Score = 173 Optimized Score = 173 Significance = 0.36
Residue Identity = 83% Matches = 173 Mismatches = 35
Gaps = 0 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAACCCCTACGACCTGGCCCGGCGGACCGAGGTTCCGGTGGAGCCGACGTCG
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
TCAAGAGAACCCCTACGACCTGGCCCGGCGGCGGACCGAGGTTCCGGTGGAGCCGACGTCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGATCCGATACGACGCTCCACGCTGACCGAAGAGCTGTCGCCACCATCGATGACCTGCTCGCTGC
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
AGAATGCCCAACCCACGACCTGACCGAAGAGAGAGCTGTCGCCACCATCGATGACCTGCTCGCTGC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGTCAGACACCGTACGACCTGGCCCGGCGGACCGAGGTTCCGGTGGAGCCGACGACAT
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
ACGAGGCGACGACGATGAGGTCCCGGCGGTGAGGTCCGGTGAAGCCGACGACAT
150 160 170 180 190 200 X

16. US-09-697-123b-3 (1-208)

US-09-697-123b-2 Sequence 20, Application US/09697123b

Initial Score = 170 Optimized Score = 182 Significance = 0.30
Residue Identity = 90% Matches = 182 Mismatches = 17
Gaps = 3 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAACCCCTACGACCTGGCCCGGTGCGCCCTACAGGTCAACAAGAGCTGGCCCTGACAGTCG
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
TCAAGAGAACCCCTACGACCTGGCCCGGTGCGCCCTACAGGTCAACAAGAGCTGGCCCTGACAGTCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGATCCGATACACCGTCCACGCTGACCGAAGAGAGCTGTCGCCACCATCGATGACCTGCTCGCTGC
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
GCGATCCGATACACCGTCCACGCTGACCGAAGAGAGCTGTCGCCACCATCGATGACCTGCTCGCTGC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGTCAGACACCGTACGACCTGGCCCGGCGGACCGAGGTTCCGGTGGAGAGCCGACGACAT
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
ACGAGGCGACCGCCACGATGACCGTCC--CGCATGAGGTTGCGGTGAGAGCCGACGACAT
150 160 170 180 190 200 X

17. US-09-697-123b-3 (1-208)

US-09-697-123b-1 Sequence 12, Application US/09697123b

Initial Score = 163 Optimized Score = 184 Significance = 0.16
Residue Identity = 90% Matches = 184 Mismatches = 18
Gaps = 1 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAACCCCTACGACCTGGCCCGGTGCGCCCTACAGGTCAACAAGAGCTGGCCCTGACAGTCG
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
TCAAGAGAACCCCTACGACCTGGCCCGGTGCGCCCTACAGGTCAACAAGAGCTGGCCCTGACAGTCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGATCCGATACACCGTCCACGCTGACCGAAGAGAGCTGTCGCCACCATCGATGACCTGCTCGCTGC
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
GCGATCCGATACACCGTCCACGCTGACCGAAGAGAGATGCTGCGACCATCGATGACCTGCTCGCTGC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGTCAGACACCGATGACCTGGCCCGGCGGACCGAGGTTCCGGTGGAGAGCCGACGACAT
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
ACGAGGCGACGACGATGACCTGGCCCGGCGGCGGACCGAGGTTCCGGTGGAGAGCCGACGACAT
150 160 170 180 190 200 X

18. US-09-697-123b-3 (1-208)

US-09-697-123b-4 Sequence 4, Application US/09697123b

Initial Score = 160 Optimized Score = 195 Significance = 0.10
Residue Identity = 96% Matches = 200 Mismatches = 7
Gaps = 1 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAACCCCTACGACCTGGCCCGGCGGCGGACCGAGGTTCCGGTGGAGAGCTGGCCCTGACAGTCG
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
TCAAGAGAACCCCTACGACCTGGCCCGGCGGCGGCGGACCGAGGTTCCGGTGGAGAGCTGGCCCTGACAGTCG
X 10 20 30 40 50 60 70

80 90 100 110 120 130 140
GCGATCCGATACACCGTCCACGCTGACCGAAGAGAGCTGTCGCCACCATCGATGACCTGCTCGCTGC
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
GCGATCCGATACACCGTCCACGCTGACCGAAGAGAGCTGTCGCCACCATCGATGACCTGCTCGCTGC
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGTCAGACACCGATGACCTGGCCCGGCGGACCGAGGTTCCGGTGGAGAGCCGACGACAT
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||
ACGAGGCGACGACGATGACCTGGCCCGGCGGCGGACCGAGGTTCCGGTGGAGAGCCGACGACAT
150 160 170 180 190 200 X

19. US-09-697-123b-3 (1-208)

US-09-697-123b-1 Sequence 13, Application US/09697123b

Initial Score = 142 Optimized Score = 167 Significance = -0.26
Residue Identity = 79% Matches = 167 Mismatches = 31
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAAGAGAACCCCTACGACCTGGCCCGGTGCGCCCTACAGGTCAACAAGAGCTGGCCCTGACAGTCG
||||| ||| ||||||||| ||||||| ||||||| ||||||| |||||||

```
|||||
TCAAGAGAAAGCCCTACGACCTGCCCCGGTGGCGCTACAAAGTCAACAGAAAGAACTGGGCTGAAACCG
X      10      20      30      40      50      60      70
      80      90      100     110     120     130     140
GGGATCCGATCACCACCTCCACGCTGACGAGAAAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
|||||
ATATCGGATCACCACGACGACCTCCACGAGAAAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
      80      90      100     110     120     130     140
ACGAGCTGACGA-----CACGATGACCGTTCCGCGCGGACCGACCGAGTTCCGCTGAGACCG
|||||
ACACGCGCTCTCAGGCTGGCGGACCGCCCGCTTATGACTGCCCCGGGGGTGCGAGTCCGCTGAAACCG
130     160     170     180     190     200     210
X
ACGACAT
|||||
ACGACAT
220 X
```

20. US-09-697-123B-3 (1-208)
US-09-697-123B-9 Sequence 9, Application US/09697123B

```
Initial Score = 133 Optimized Score = 171 Significance = -0.43
Residue Identity = 82% Matches = 185 Mismatches = 23
Gaps = 15 Conservative Substitutions = 3
X      10      20      30      40      50      60      70
TCAAGAGAAAGCCCTACGACCTGACGAGAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
|||||
TCAAGAGAAAGCCCTACGACCTGACGAGAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
X      10      20      30      40      50      60      70
GGCATCCGATCACCACCTCCACGCTGACGAGAAAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
|||||
GCGAGCGGATCAGCTGCTCCACGCTGACGAGAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
      80      90      100     110     120     130     140
AC-----GAGGGTCAACACAC--GATGACCGCTTCCGGGCGGACCGAGGTTCCGCTGAGACCG
|||||
ACACGCGCTTACGAGATGGCGACCGCGCTCATGACTGTCCCGGGGATCGAGGTGCGGTTGAGACCG
150     160     170     180     190     200     210
X
ACGACAT
|||||
ACGACAT
220 X
```

21. US-09-697-123B-3 (1-208)
US-09-697-123B-7 Sequence 7, Application US/09697123B

```
Initial Score = 125 Optimized Score = 164 Significance = -0.59
Residue Identity = 79% Matches = 171 Mismatches = 37
Gaps = 6 Conservative Substitutions = 0
X      10      20      30      40      50      60      70
TCAAGAGAAAGCCCTACGACCTGACGAGAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
|||||
TCAAGAGAAAGCCCTACGACCTGACGAGAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
X      10      20      30      40      50      60      70
GC-----GATTCGATCACCAGCTCCACGCTGACGAGAGAGAGCTGCGCACATCGAATGATCTGCTGC
|||||
CCAAACCGGCTTGTGTAATGCGCACCGCTCACCGGAGAGAGAGCTGCGCACATCGAATGATCTGCTGC
      80      90      100     110     120     130     140
```

```
140     150     160     170     180     190     200
GCTTCACGAGAGGCTGACGACGATGACCTTCCGGGCGGACCGAGGTTCCGCTGAGACCGACGACAT
|||||
GCTTCACGAGAGGCTGACGACGATGACCTTCCGGGCGGACCGAGGTTCCGCTGAGAGTGCAGCAT
150     160     170     180     190     200     210 X
```

22. US-09-697-123B-3 (1-208)
US-09-697-123B-1 Sequence 18, Application US/09697123B

```
Initial Score = 112 Optimized Score = 160 Significance = -0.85
Residue Identity = 78% Matches = 166 Mismatches = 42
Gaps = 3 Conservative Substitutions = 0
X      10      20      30      40      50      60      70
TCAAGAGAAAGCCCTACGACCTGACGAGAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
|||||
TCAAGAGAAAGCCCTACGACCTGACGAGAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
X      10      20      30      40      50      60      70
```

```
80      90      100     110     120     130     140
GGCATCCGATCACCACGCT---CCACGCTGACCGAGAGAGAGCTGCGCACATCGAATGATCTGCTGC
|||||
ACCGCGCGACGACGACCTCGACGACCTGACCGAGAGAGAGCTGCTGCGCACATCGAATGATCTGCTGC
      80      90      100     110     120     130     140
```

```
150     160     170     180     190     200
TCCAGAGAGAGGCTGACGACGATGACCTTCCGGGCGGACCGAGGTTCCGCTGAGAGACGACGAT
|||||
TCCATCAGAGGCGCAAGAGAGATGACCTTCCGGGCGGAGTGCAGAGTCCGCTGAGAGTGCAGAGAT
150     160     170     180     190     200     210
```

23. US-09-697-123B-3 (1-208)
US-09-697-123B-2 Sequence 21, Application US/09697123B

```
Initial Score = 112 Optimized Score = 166 Significance = -0.85
Residue Identity = 80% Matches = 173 Mismatches = 35
Gaps = 6 Conservative Substitutions = 0
X      10      20      30      40      50      60      70
TCAAGAGAAAGCCCTACGACCTGACGAGAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
|||||
TCAAGAGAAAGCCCTACGACCTGACGAGAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
X      10      20      30      40      50      60      70
```

```
80      90      100     110     120     130
GGCATCC-----GATTCACCCACGCTCCACGCTGACCGAGAGAGAGCTGCGCACATCGAATGATCTGCTGC
|||||
CCAAATCCGCTGAGGTGACACCCACGACCTTCCACCGAGAGAGAGCTGCGCACATCGAATGATCTGCTGC
      80      90      100     110     120     130     140
```

```
140     150     160     170     180     190     200
GCTTCGACGAGAGGCTGACGACGATGACCTTCCGGGCGGACCGAGGTTCCGCTGAGAGACCGACGAT
|||||
GCTTCGACGAGAGGCTGACGACGATGACCTTCCGGGCGGAGTGTGAGAGAGAT
150     160     170     180     190     200     210 X
```

24. US-09-697-123B-3 (1-208)
US-09-697-123B-1 Sequence 14, Application US/09697123B

```
Initial Score = 112 Optimized Score = 168 Significance = -0.85
Residue Identity = 81% Matches = 175 Mismatches = 33
Gaps = 6 Conservative Substitutions = 0
X      10      20      30      40      50      60      70
TCAAGAGAAAGCCCTACGACCTGACGAGAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
|||||
TCAAGAGAAAGCCCTACGACCTGACGAGAGAGAGCTGCGCACATCGAATGATCTGCTGCTGC
X      10      20      30      40      50      60      70
```

```
      80      90      100      110      120      130
CGGATCCG-----ATCACCAGCTCCAGCGTGCAGGAAAGACAGCTCCGCGCCACCATCGAGTACCTGCTCC
|||||
CCGAGTCGGCCGTACCCGCTCGACCGACGCTGACCGAAGCGGATGTCGTCCGACCATCGAGTACCTGCTGC
      80      90      100      110      120      130      140
140      150      160      170      180      190      200
GTCGTCCAGCAGGGGTGACGACGATGATGACCGCTCCGGGGGCGACCGAGGTTCCGGTGGAGACCGAGACAT
|||||
GCCTGCACGAGGGCCAGCGCAACGATGACGCTTCCGCGCGCGTGCAGGTTGCCGTTGGAGACCGAGACAT
150      160      170      180      190      200      210
X
```

25. US-09-697-123B-3 (1-208)
US-09-697-123B-2 Sequence 25, Application US/09697123B

Initial Score	=	19	Optimized Score	=	19	Significance	=	-2.69
Residue Identity	=	100%	Matches	=	19	Mismatches	=	0
Gaps	=	0	Conservative Substitutions	=			=	0

X	10	X
TCAGGAGAGCGCTACGA		
TCAGGAGAGCGCTACGA		
X	10	X

Sequence Name	Description	Init. length	Opt. Score	Frame
---------------	-------------	--------------	------------	-------

[illegible]

X
10
20
30
40
50
60
70
TCAAGGAGAGCGCTACGACCTGGCCCCGGGTGGCCGCTACAAGGTCACAAAGAAGCTCGGTGTGAACCTCG

|||||
TCAGAGAGAGGCTACGACCTGGGCGCGCTGCGCTACAGGTCAACAGAAAGCTGGCTGACGCGC
X 10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCAAGCCGATTCACAGCTGACGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
|||||
GCGAGCCGATTCACGCTGACGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
80 90 100 110 120 130 140
ACGAGGCTGACGCTGACGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
150 160 170 180 190 200 210
X
ACGACAT
|||||
ACGACAT
220 X

20. US-09-697-123B-2 (1-208)
US-09-697-123B-1 Sequence 13, Application US/09697123B

Initial Score = 140 Optimized Score = 166 Significance = -0.30
Residue Identity = 78% Matches = 176 Mismatches = 32
Gaps = 15 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
|||||
TCAGAGAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
X 10 20 30 40 50 60 70
GCAAGCCGATTCACAGCTGACGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
|||||
ATCATGCGATTCACAGCTGACGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
80 90 100 110 120 130 140
ACGAGGCTGACGCTGACGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
150 160 170 180 190 200 210
X
ACGACAT
|||||
ACGACAT
220 X

21. US-09-697-123B-2 (1-208)
US-09-697-123B-7 Sequence 7, Application US/09697123B

Initial Score = 129 Optimized Score = 166 Significance = -0.52
Residue Identity = 80% Matches = 173 Mismatches = 35
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
|||||
TCAGAGAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
X 10 20 30 40 50 60 70
GCAAGCCGATTCACAGCTGACGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
|||||
GCAAGCCGATTCACAGCTGACGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
80 90 100 110 120 130 140

140 150 160 170 180 190 200
GCGTCACAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
|||||
GCGTCACAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
150 160 170 180 190 200 210 X

22. US-09-697-123B-2 (1-208)
US-09-697-123B-1 Sequence 18, Application US/09697123B

Initial Score = 117 Optimized Score = 161 Significance = -0.76
Residue Identity = 79% Matches = 167 Mismatches = 41
Gaps = 3 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
|||||
TCAGAGAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
X 10 20 30 40 50 60 70
GCAAGCCGATTCACAGCTGACGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
|||||
ACCGCGCGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
80 90 100 110 120 130 140
TGACAGAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
150 160 170 180 190 200 210
TGACAGAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
150 160 170 180 190 200 210

23. US-09-697-123B-2 (1-208)
US-09-697-123B-2 Sequence 21, Application US/09697123B

Initial Score = 112 Optimized Score = 167 Significance = -0.86
Residue Identity = 81% Matches = 174 Mismatches = 34
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
|||||
TCAGAGAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
X 10 20 30 40 50 60 70
GCAAGCCGATTCACAGCTGACGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
|||||
GCAAGCCGATTCACAGCTGACGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
80 90 100 110 120 130 140
GCGTCACAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
150 160 170 180 190 200 210 X
GCGTCACAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
150 160 170 180 190 200 210 X

24. US-09-697-123B-2 (1-208)
US-09-697-123B-1 Sequence 14, Application US/09697123B

Initial Score = 112 Optimized Score = 170 Significance = -0.86
Residue Identity = 82% Matches = 177 Mismatches = 31
Gaps = 6 Conservative Substitutions = 0

X 10 20 30 40 50 60 70
TCAGAGAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
|||||
TCAGAGAGAGGCTACGACCTGGGCGCGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
X 10 20 30 40 50 60 70
GCAAGCCGATTCACAGCTGACGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
|||||
GCAAGCCGATTCACAGCTGACGCTGACGAGAAAGAGCTGATGACCATGACGCTGGGCGCGC
80 90 100 110 120 130 140


```
Initial Score      - 183  Optimized Score      - 183  Significance      - 0.52
Residue Identity  - 87%  Matches              - 183  Mismatches        - 25
```

Gaps - 0 Conservative Substitutions - 0

10. US-09-697-123B-1 (1-208)
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.50
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X
TCAAGGAGAGGGCTACGACCTGGCGGCGGTACAGCCCTCAACAAGAGCTCGCCCTGCACGCTGC
|||||
TCAAGGAGAGGGCTACGACCTGGCGGCGGTACAGCCCTCAACAAGAGCTCGCCCTGCACGACCA
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCCATCCGATACACGACCTCCAGCTGACCGAGGAAGAGCTGTCGCCACCATTCAGATCTGGTCCGCTGC
|||||
ATCATCCGATACACGACCTCCAGCTGACCGAGGAAGAGCTGTCGCCACCATTCAGATCTGGTCCGCTGC
80 90 100 110 120 130 140
150 160 170 180 190 200
ACGAGGGCCAGCAGACGATGACCTGCGGCGGCGGTGAGAGTGCCTGAGAGCTGAGAGCTGAGAGCTG
|||||
ACGAGGGCCAGCAGCAGATGACCTGCGGCGGCGGTGAGAGTGCCTGAGAGCTGAGAGCTGAGAGCTG
150 160 170 180 190 200
X

9. US-09-697-123B-1 (1-208)
US-09-697-123B-1 Sequence 17, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.50
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X
TCAAGGAGAGGGCTACGACCTGGCGGCGGTACAGCCCTCAACAAGAGCTCGCCCTGCACGCTGC
|||||
TCAAGGAGAGGGCTACGACCTGGCGGCGGTACAGCCCTCAACAAGAGCTCGCCCTGCACGACCG
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCCATCCGATACACGACCTCCAGCTGACCGAGGAAGAGCTGTCGCCACCATTCAGATCTGGTCCGCTGC
|||||
CGTCCCGATACACGACCTCCAGCTGACCGAGGAAGAGCTGTCGCCACCATTCAGATCTGGTCCGCTGC
80 90 100 110 120 130 140
150 160 170 180 190 200
ACGAGGGCCAGCAGACGATGACCTGCGGCGGCGGTGAGAGTGCCTGAGAGCTGAGAGCTGAGAGCTG
|||||
ACGAGGGCCAGCAGCAGATGACCTGCGGCGGCGGTGAGAGTGCCTGAGAGCTGAGAGCTGAGAGCTG
150 160 170 180 190 200
X

10. US-09-697-123B-1 (1-208)
US-09-697-123B-1 Sequence 16, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.50
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X
TCAAGGAGAGGGCTACGACCTGGCGGCGGTACAGCCCTCAACAAGAGCTCGCCCTGCACGCTGC
|||||
TCAAGGAGAGGGCTACGACCTGGCGGCGGTACAGCCCTCAACAAGAGCTCGCCCTGCACGCTGC
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCCATCCGATACACGACCTCCAGCTGACCGAGGAAGAGCTGTCGCCACCATTCAGATCTGGTCCGCTGC
|||||
GCGAGCCCATTCACGACCTCCAGCTGACCGAGGAAGAGCTGTCGCCACCATTCAGATCTGGTCCGCTGC
80 90 100 110 120 130 140
150 160 170 180 190 200
ACGAGGGCCAGCAGACGATGACCTGCGGCGGCGGTGAGAGTGCCTGAGAGCTGAGAGCTGAGAGCTG
|||||
ACGAGGGCCAGCAGCAGATGACCTGCGGCGGCGGTGAGAGTGCCTGAGAGCTGAGAGCTGAGAGCTG
150 160 170 180 190 200
X

11. US-09-697-123B-1 (1-208)
US-09-697-123B-6 Sequence 6, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.50
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X
TCAAGGAGAGGGCTACGACCTGGCGGCGGTACAGCCCTCAACAAGAGCTCGCCCTGCACGCTGC
|||||
TCAAGGAGAGGGCTACGACCTGGCGGCGGTACAGCCCTCAACAAGAGCTCGCCCTGCACGCTGC
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCCATCCGATACACGACCTCCAGCTGACCGAGGAAGAGCTGTCGCCACCATTCAGATCTGGTCCGCTGC
|||||
GCGAGCCCATTCACGACCTCCAGCTGACCGAGGAAGAGCTGTCGCCACCATTCAGATCTGGTCCGCTGC
80 90 100 110 120 130 140
150 160 170 180 190 200
ACGAGGGCCAGCAGACGATGACCTGCGGCGGCGGTGAGAGTGCCTGAGAGCTGAGAGCTGAGAGCTG
|||||
ACGAGGGCCAGCAGCAGATGACCTGCGGCGGCGGTGAGAGTGCCTGAGAGCTGAGAGCTGAGAGCTG
150 160 170 180 190 200
X

12. US-09-697-123B-1 (1-208)
US-09-697-123B-5 Sequence 5, Application US/09697123B

Initial Score = 182 Optimized Score = 182 Significance = 0.50
Residue Identity = 87% Matches = 182 Mismatches = 26
Gaps = 0 Conservative Substitutions = 0

X
TCAAGGAGAGGGCTACGACCTGGCGGCGGTACAGCCCTCAACAAGAGCTCGCCCTGCACGCTGC
|||||
TCAAGGAGAGGGCTACGACCTGGCGGCGGTACAGCCCTCAACAAGAGCTCGCCCTGCACGCTGC
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCCATCCGATACACGACCTCCAGCTGACCGAGGAAGAGCTGTCGCCACCATTCAGATCTGGTCCGCTGC
|||||
GCGAGCCCATTCACGACCTCCAGCTGACCGAGGAAGAGCTGTCGCCACCATTCAGATCTGGTCCGCTGC
80 90 100 110 120 130 140
150 160 170 180 190 200
ACGAGGGCCAGCAGACGATGACCTGCGGCGGCGGTGAGAGTGCCTGAGAGCTGAGAGCTGAGAGCTG
|||||
ACGAGGGCCAGCAGCAGATGACCTGCGGCGGCGGTGAGAGTGCCTGAGAGCTGAGAGCTGAGAGCTG
150 160 170 180 190 200
X

13. US-09-697-123B-1 (1-208)
US-09-697-123B-2 Sequence 23, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.48
Residue Identity = 87% Matches = 181 Mismatches = 27
Gaps = 0 Conservative Substitutions = 0

X
TCAAGGAGAGGGCTACGACCTGGCGGCGGTACAGCCCTCAACAAGAGCTCGCCCTGCACGCTGC
|||||
TCAAGGAGAGGGCTACGACCTGGCGGCGGTACAGCCCTCAACAAGAGCTCGCCCTGCACGCTGC
X
10 20 30 40 50 60 70
80 90 100 110 120 130 140
GCCATCCGATACACGACCTCCAGCTGACCGAGGAAGAGCTGTCGCCACCATTCAGATCTGGTCCGCTGC
|||||
GCGAGCCCATTCACGACCTCCAGCTGACCGAGGAAGAGCTGTCGCCACCATTCAGATCTGGTCCGCTGC
80 90 100 110 120 130 140
150 160 170 180 190 200
ACGAGGGCCAGCAGACGATGACCTGCGGCGGCGGTGAGAGTGCCTGAGAGCTGAGAGCTGAGAGCTG
|||||
ACGAGGGCCAGCAGCAGATGACCTGCGGCGGCGGTGAGAGTGCCTGAGAGCTGAGAGCTGAGAGCTG
150 160 170 180 190 200
X

|||||
ATGACGGATCATTCCAGCCTTCCAGGTGCGAGGTGCTGAGTACGACGACAT
150 160 170 180 190 200 X

14. US-09-697-123B-1 (1-208)
US-09-697-123B-2 Sequence 22, Application US/09697123B

Initial Score = 181 Optimized Score = 181 Significance = 0.48
Residue Identity = 87% Matches = 181 Mismatches = 27
Gaps = 0 Conservative Substitutions = 0

X
TCAAGAGAACGGCTACGACCTGCGCCGGCTAGCGCCGCTACAAAGTCAACAAAGAGCTGCGCTCAGCTCG
|||||
TCAAGAGAACGGCTACGACCTGCGCCGGCTAGCGCCGCTACAAAGTCAACAAAGAGCTGCGCTCAGCTCG
150 160 170 180 190 200 X

80 90 100 110 120 130 140
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGCGTCTGCGCCAGCATCGAGTACCTGCTGCGCTCG
|||||
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGCGTCTGCGCCAGCATCGAGTACCTGCTGCGCTCG
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCCAGACACAGATGACCTGCGCCGGCGACACGAGGTGCGCGTTGAGACCGACGACAT
|||||
ACGAGGCCAGACACAGATGACCTGCGCCGGCGACACGAGGTGCGCGTTGAGACCGACGACAT
150 160 170 180 190 200 X

15. US-09-697-123B-1 (1-208)
US-09-697-123B-2 Sequence 24, Application US/09697123B

Initial Score = 175 Optimized Score = 175 Significance = 0.36
Residue Identity = 84% Matches = 175 Mismatches = 33
Gaps = 0 Conservative Substitutions = 0

X
TCAAGAGAACGGCTACGACCTGCGCCGGCGACACGAGGTGCGCGTTGAGACCGACGACAT
|||||
TCAAGAGAACGGCTACGACCTGCGCCGGCGACACGAGGTGCGCGTTGAGACCGACGACAT
150 160 170 180 190 200 X

80 90 100 110 120 130 140
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGCTGCTGCGCCATCGAGTACCTGCTGCGCTCG
|||||
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGCTGCTGCGCCATCGAGTACCTGCTGCGCTCG
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCCAGACACAGATGACCTGCGCCGGCGACACGAGGTGCGCGTTGAGACCGACGACAT
|||||
ACGAGGCCAGACACAGATGACCTGCGCCGGCGACACGAGGTGCGCGTTGAGACCGACGACAT
150 160 170 180 190 200 X

16. US-09-697-123B-1 (1-208)
US-09-697-123B-2 Sequence 20, Application US/09697123B

Initial Score = 172 Optimized Score = 184 Significance = 0.30
Residue Identity = 91% Matches = 190 Mismatches = 15
Gaps = 3 Conservative Substitutions = 0

X
TCAAGAGAACGGCTACGACCTGCGCCGGCGTACGAGGTCAACAAAGAGCTGCGCTCAGCTCG
|||||
TCAAGAGAACGGCTACGACCTGCGCCGGCGTACGAGGTCAACAAAGAGCTGCGCTCAGCTCG
150 160 170 180 190 200 X

80 90 100 110 120 130 140
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGCTGCTGCGCCAGCATCGAGTACCTGCTGCGCTCG
|||||
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGCTGCTGCGCCAGCATCGAGTACCTGCTGCGCTCG
80 90 100 110 120 130 140

GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGCTGCTGCGCCAGCATCGAGTACCTGCTGCGCTCG
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCCAGACACAGATGACCTGCGCCGGCGACACGAGGTGCGCGTTGAGACCGACGACAT
|||||
ACGAGGCCAGACACAGATGACCTGCGCCGGCGACACGAGGTGCGCGTTGAGACCGACGACAT
150 160 170 180 190 200 X

17. US-09-697-123B-1 (1-208)
US-09-697-123B-4 Sequence 4, Application US/09697123B

Initial Score = 162 Optimized Score = 189 Significance = 0.10
Residue Identity = 93% Matches = 194 Mismatches = 13
Gaps = 1 Conservative Substitutions = 0

X
TCAAGAGAACGGCTACGACCTGCGCCGGCGTACGAGGTCAACAAAGAGCTGCGCTCAGCTCG
|||||
TCAAGAGAACGGCTACGACCTGCGCCGGCGTACGAGGTCAACAAAGAGCTGCGCTCAGCTCG
150 160 170 180 190 200 X

80 90 100 110 120 130 140
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGCTGCTGCGCCATCGAGTACCTGCTGCGCTCG
|||||
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGCTGCTGCGCCATCGAGTACCTGCTGCGCTCG
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCCAGACACAGATGACCTGCGCCGGCGACACGAGGTGCGCGTTGAGACCGACGACAT
|||||
ACGAGGCCAGACACAGATGACCTGCGCCGGCGACACGAGGTGCGCGTTGAGACCGACGACAT
150 160 170 180 190 200 X

18. US-09-697-123B-1 (1-208)
US-09-697-123B-1 Sequence 12, Application US/09697123B

Initial Score = 159 Optimized Score = 182 Significance = 0.04
Residue Identity = 89% Matches = 187 Mismatches = 20
Gaps = 1 Conservative Substitutions = 0

X
TCAAGAGAACGGCTACGACCTGCGCCGGCGTACGAGGTCAACAAAGAGCTGCGCTCAGCTCG
|||||
TCAAGAGAACGGCTACGACCTGCGCCGGCGTACGAGGTCAACAAAGAGCTGCGCTCAGCTCG
150 160 170 180 190 200 X

80 90 100 110 120 130 140
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGCTGCTGCGCCATCGAGTACCTGCTGCGCTCG
|||||
GCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGCTGCTGCGCCATCGAGTACCTGCTGCGCTCG
80 90 100 110 120 130 140

150 160 170 180 190 200 X
ACGAGGCCAGACACAGATGACCTGCGCCGGCGACACGAGGTGCGCGTTGAGACCGACGACAT
|||||
ACGAGGCCAGACACAGATGACCTGCGCCGGCGACACGAGGTGCGCGTTGAGACCGACGACAT
150 160 170 180 190 200 X

19. US-09-697-123B-1 (1-208)
US-09-697-123B-9 Sequence 9, Application US/09697123B

Initial Score = 146 Optimized Score = 173 Significance = -0.22
Residue Identity = 83% Matches = 187 Mismatches = 21
Gaps = 15 Conservative Substitutions = 0

X
TCAAGAGAACGGCTACGACCTGCGCCGGCGTACGAGGTCAACAAAGAGCTGCGCTCAGCTCG
|||||
TCAAGAGAACGGCTACGACCTGCGCCGGCGTACGAGGTCAACAAAGAGCTGCGCTCAGCTCG
150 160 170 180 190 200 X

[illegible]

Initial Score	-	109	Optimized Score	-	170	Significance	-	-0.96
Residue Identity	-	82%	Matches	-	177	Mismatches	-	31
Gaps	-	6	Conservative Substitutions	-			-	0

GGGATCC-----GATCACCAGCTCCACGCTGACCGAGGAGAACCTGCTGCCACCATCGACTACCTGCTCC

140 GCTCGCAGCGAGCCGACGACGATGATCCCGGGGGGCGACCGAGCGCTTGAGACCGACGACAT X
 150 GGGGTT
 160 GCGCGCAGCGAGGCCGACGACGATGATCCCGGGGGGCGCTTGAGCTGCTCGGTGAGTGTGGAGCAAT
 170
 180
 190
 200
 210 X

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 : Search time 106.647 Seconds

(Without alignments)
3037.202 Million cell updates/sec

Title: US-09-697-123B-26

Perfect score: 1 ggaatgtgacagcagctctgc 20

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estda: *
2: em_esthum: *
3: em_estlin: *
4: em_estnu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hic: *
9: gb_est1: *
10: gb_est2: *
11: gb_hic: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gsa: *
18: em_gsa_hum: *
19: em_gsa_inv: *
20: em_gsa_pin: *
21: em_gsa_vrt: *
22: em_gsa_fun: *
23: em_gsa_mam: *
24: em_gsa_mus: *
25: em_gsa_other: *
26: em_gsa_pro: *
27: em_gsa_tod: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	95.0	414	17	A2606523 1M0428B22
2	18.4	92.0	377	9	AA143959 ms13d11.r
3	18.4	92.0	437	10	AA211046 uo90e08.y
4	18.4	92.0	502	9	AL045401 DKFZP434D
5	18.4	92.0	579	9	AI603914 ms13d11.y
6	18.4	92.0	585	13	BI697763 603346856

C	7	18.4	92.0	597	14	B0554372
C	8	18.4	92.0	638	10	AN272508
C	9	18.4	92.0	706	14	BM943312
C	10	18.4	92.0	738	13	BI697529
C	11	18.4	92.0	759	12	BC985603
C	12	18.4	92.0	768	13	BI688003
C	13	18.4	92.0	778	13	BI695451
C	14	18.4	92.0	824	13	BI155204
C	15	18.4	92.0	4072	11	BC028426
C	16	18.4	90.0	314	17	A2276243
C	17	17.4	87.0	279	9	AA144678
C	18	17.4	87.0	532	17	A0732019
C	19	17	85.0	306	10	AM114205
C	20	17	85.0	950	17	CNS0442C
C	21	16.8	84.0	271	17	CNS024LB
C	22	16.8	84.0	272	10	BB231701
C	23	16.8	84.0	380	12	BC303432
C	24	16.8	84.0	417	13	BI297125
C	25	16.8	84.0	461	10	BI866562
C	26	16.8	84.0	477	12	BF554685
C	27	16.8	84.0	530	13	BI843127
C	28	16.8	84.0	538	10	AM175549
C	29	16.8	84.0	550	13	BI671235
C	30	16.8	84.0	554	17	BH035861
C	31	16.8	84.0	641	14	BQ904110
C	32	16.8	84.0	652	10	AM915754
C	33	16.8	84.0	861	13	BI193716
C	34	16.4	82.0	302	13	BI049959
C	35	16.4	82.0	459	9	A1260450
C	36	16.4	82.0	520	17	A2796561
C	37	16.4	82.0	521	9	AA698485
C	38	16.4	82.0	538	13	BM405728
C	39	16.4	82.0	571	17	AO668780
C	40	16.4	82.0	602	12	BF153712
C	41	16.4	82.0	627	17	AO757049
C	42	16.4	82.0	680	12	BC590263
C	43	16.4	82.0	703	13	BI767119
C	44	16.4	82.0	732	12	BC886620
C	45	16.4	82.0	776	13	BM408140

ALIGNMENTS

B0554372 H4027G03-
AN272508 xu20b08.x
BM943312 UT-M-ER0P
BI697529 603349086
BC985603 602453904
BI688003 603315768
BI695451 603346455
BI155204 602903188
BC028426 Homo_sapi
A2276243 RPECI-23-1
AA144678 mr69c07.r
A0732019 HS_5555_A
AM114205 rs70a01.y
AL314108 Tetradon
AL180920 Tetradon
BB231701 BB231701
BC303432 f15f03.x
BI297125 UT-R-DK0-
BI866562 f15f05.y
BF554685 UT-R-C0-1
BI843127 f15f05.x
AM175549 f13632.x
BI671235 f08210.x
BH035861 RPECI-24-3
BQ904110 Ta03_2090
AM915754 EST347058
BI193716 602946680
BI049959 CM2-GM029
A1260450 LP04321.5
A2796561 2M0052A23
AA698485 HL04775.5
BM405728 EST580055
AO668780 HS_5435_A
BF153712 046F08_Ms
AO757049 HS_5448_B
BC590263 EST498093
BI767119 603054180
BC886620 EST512471
BM408140 EST582467

RESULT 1
LOCUS 414 bp DNA linear GSS 13-DEC-2000
DEFINITION 1M0428B22F Mouse 10kb plasmid U0CC1M library Mus musculus genomic
clone U0CC1M0428B22 F, DNA sequence.
ACCESSION A2606523
VERSION A2606523.1 GI:11728629
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 414)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
CONTACT Contact: Robert B. Weiss
UNIVERSITY University of Utah
INSTITUTION University of Utah Genome Center
ADDRESS Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
TEL: 801 585 5606
FAX: 801 585 7177
EMAIL: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00

FEATURES

Plate: 0428 row: B column: 22
Seq primer: CCTGTAAACGACGCCACG
Class: plasmid ends
High quality sequence stop: 414.
Location/Qualifiers

1..414

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUCGCM0428B2"

/clone_lib="Mouse 10kb plasmid UUCGCM library"

/sex="Male"

/note="Vector: E. coli strain XL10-Gold, T1-resistant, F-"

/musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pMD42 (g147321149b1AF12972.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT

81 a 100 c 87 g 146 t

ORIGIN

Query Match 95.0%; Score 19; DB 17; Length 414;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATGTTGATCAGGCTGTC 20
|||||

Db 383 GATGTTGATCAGGCTGTC 401

RESULT 2

AA143959/c

LOCUS

DEFINITION

AA143959 377 bp mRNA linear EST 12-FEB-1997
m13d11.r1 Stratiogene mouse skin (4937313) Mus musculus cDNA clone
IMAGE:606837 5 similar to TR:G498013 G498013 X104. ; mRNA

SEQUENCE

AA143959.1 GI:1713346

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu

This clone is available royalty-free through LNL; contact the

FEATURES

source

IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:372269

Possible reversed clone: similarity on wrong strand

Seq primer: -28m13 rev1 ET from Amersham

High quality sequence stop: 331.
Location/Qualifiers

1..377

/organism="Mus musculus"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:606837"

/clone_lib="Stratiogene mouse skin (#937313)"

/sex="females"

/tissue_type="whole skin"

/dev_stage="11 weeks old"

/note="Organ: skin; Vector: pBluescript SK-; Site.1: EcoRI

dr. whole skin from 11 week old C57BL/6 female mice.

Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'

adaptor sequence: 5' GAATTCGACGAG 3' -3' adaptor

sequence: 5' CTCGAGTTT TTTT TTTT TTTT 3'

104 a 107 c 102 g 64 t

BASE COUNT

ORIGIN

Query Match 92.0%; Score 18.4; DB 9; Length 377;
Best Local Similarity 95.0%; Pred. No. 2e+02;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCATGTTGATCAGGCTGTC 20
|||||

Db 54 GCATGTTGATCAGGCTGTC 35

RESULT 3

AA211046/c

LOCUS

DEFINITION

AA211046 437 bp mRNA linear EST 03-DEC-1999
uo90e08.y1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:2649830 5'
similar to TR:Q92001 Q92001 TIGHT JUNCTION PROTEIN ZO-2. ; mRNA

SEQUENCE

AA211046.1 GI:6516986

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

```

/dev-stage="10 months"
/lab-host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: Saliv;
Site:2: Notti; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
BASE COUNT      136 a   102 c   111 g   86 t   2 others
ORIGIN

Query Match      92.0%; Score 18.4; DB 10; Length 437;
Best Local Similarity 95.0%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 GGATGTTGATCAGGCTCTGC 20
|||||  |||||  |||||  |||||  |||||
Db      306 GGATGTTGATCAGGCTCTGC 287

RESULT 4
AL045401/c      502 bp      mRNA      linear      EST 29-FEB-2000
LOCUS           DKF2P34D055.r1 434 (synonym: htes3) Homo sapiens cDNA clone
DEFINITION      DKF2P34D055 5', mRNA sequence.
ACCESSION       AL045401
VERSION         AL045401.1 GI:5433550
KEYWORDS        EST.
SOURCE          human.
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE       1 (bases 1 to 502)
Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
,S.
EST (Duesterhoeft, et al.)
JOURNAL         Unpublished (1999)
COMMENT         Contact: Duesterhoeft A
MIPS
Am Klopferplatz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
consortium of Qiagen (Hilden/Germany) within the cDNA sequencing
No sl sequence available.
This clone (DKF2P34D055) is available at the RZPD in Berlin.
Please contact the RZPD: Reesourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..502
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2P34D055"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site:1: Notti; Site:2: Saliv"
BASE COUNT      152 a   123 c   122 g   101 t   4 others
ORIGIN

Query Match      92.0%; Score 18.4; DB 9; Length 502;
Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 GGATGTTGATCAGGCTCTGC 20
|||||  |||||  |||||  |||||  |||||
Db      30 GGATGTTGATCAGGCTCTGC 11

RESULT 5
AL0603914/c

```

```

LOCUS           A1603914      579 bp      mRNA      linear      EST 15-MAR-2000
DEFINITION      ms3d11.y1 Stratiagene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:606837 5' similar to TR:P70625 P70625 ZONULA OCCUDENS 2
ACCESSION       A1603914
VERSION         A1603914.1 GI:4613076
KEYWORDS        EST.
SOURCE          house mouse.
ORGANISM        Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
REFERENCE       1 (bases 1 to 579)
AUTHORS        Marra,M., Hillier,L., Kucaba,T., Martin,D., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McEann,R.,
Waterson,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999
JOURNAL         Unpublished (1999)
COMMENT         Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouse@waterson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the
correct orientation)
Possible reversed clone: similarity on wrong strand
NCI:372269
Seq primer: -40RP from Gibco
High quality sequence stop: 406
POLYA-No.
FEATURES
source
1..579
/location/Qualifiers
1..579
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:606837"
/clone_lib="Stratiagene mouse skin (#937313)"
/sex="females"
/tissue_type="whole skin"
/dev_stage="11 weeks old"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: skin; Vector: pBluescript SK-; Site:1: EcorI
; Site:2: XhoI; Cloned unidirectionally. Primer: Oligo
dt. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb. Uni-ZAP XR Vector: -5'
adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor
sequence: 5' CTCGAGTCTTTTCTTTTCTTTTCTTTT 3'"
BASE COUNT      160 a   146 c   170 g   101 t   2 others
ORIGIN

Query Match      92.0%; Score 18.4; DB 9; Length 579;
Best Local Similarity 95.0%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 GGATGTTGATCAGGCTCTGC 20
|||||  |||||  |||||  |||||  |||||
Db      55 GGATGTTGATCAGGCTCTGC 36

RESULT 6
BI697763/c
LOCUS           BI697763      585 bp      mRNA      linear      EST 18-SEP-2001
DEFINITION      603346856F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5374693 5',
mRNA sequence.
ACCESSION       BI697763
VERSION         BI697763.1 GI:15660392
KEYWORDS        EST.
SOURCE          house mouse.

```

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 585)
 NIH-MGC <http://mhc.mcl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
<http://image.llnl.gov>
 Plate: LLM11953 row: 0 column: 14
 High quality sequence stop: 584.
 Location/Qualifiers
 1..585
 /organism="Mus musculus"
 /strain="FVB/N-3"
 /db_xref="taxon:10090"
 /clone_image="5374693"
 /clone_lib="NCI-CGAP_Man2"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 162 a 147 c 173 g 102 t 1 others
 ORIGIN

Query Match 92.0%; Score 18.4; DB 13; Length 585;
 Best Local Similarity 95.0%; Pred. No. 2.4e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GGATGTCATCAGGCTCTGC 20
 ||||||||| |||||||||
 Db 93 GGATGTCATCAGGCTCTGC 74

RESULT 7
 B0554372 597 bp mRNA linear EST 20-JUN-2002
 LOCUS
 DEFINITION H4027G03-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone
 ACCESSION B0554372
 VERSION B0554372.1 GI:21455260
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 597)
 VanBuren, V., Plao, Y., Dudekula, D.B., Qian, Y., Carter, M.G., Martin
 , P.R., Stagg, C.A., Bassey, U., Alba, K., Hamatani, T., Kargul, G.J.,
 Luo, A.G. and Ko, M.S.H.
 Assembly, verification, and initial annotation of NIA 7.4K mouse
 cDNA clone set
 Unpublished (2002)
 Other-ESTs: H4027G03-3
 Contact: Yong Qian
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
 Email: cdna@sun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit http://lgsun.grc.nia.nih.gov/cDNA/NIA_7_4K.html for details.
 Plate: H4027 row: G column: 03
 Seq primer: -21M13 Reverse

JOURNAL
 COMMENT

High quality sequence stop: 597
 POLYA-No.
 Location/Qualifiers
 1..597
 /organism="Mus musculus"
 /strain="C57BL/6"
 /db_xref="nlaEST:H4027G03-5"
 /db_xref="taxon:10090"
 /clone_image="H4027G03"
 /clone_lib="NIA Mouse 7.4K cDNA Clone Set"
 /sex="mixed"
 /dev_stage="mixed"
 /lab_host="DH10B"
 /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
 clone is among a rearranged set of 7,407 clones from more
 than 20 cDNA libraries."

BASE COUNT 171 a 151 c 164 g 111 t
 ORIGIN

Query Match 92.0%; Score 18.4; DB 14; Length 597;
 Best Local Similarity 95.0%; Pred. No. 2.4e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GGATGTCATCAGGCTCTGC 20
 ||||||||| |||||||||
 Db 174 GGATGTCATCAGGCTCTGC 155

RESULT 8
 AM272508 638 bp mRNA linear EST 03-JAN-2000
 LOCUS
 DEFINITION XU20508.x1 NCI-CGAP-Col4 Homo sapiens cDNA clone IMAGE:2800695 3'
 similar to TR:Q15883 Q15883 X104 ;, mRNA sequence.
 ACCESSION AM272508
 VERSION AM272508.1 GI:6659538
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 638)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

JOURNAL
 COMMENT

Possible reversed clone; similarity on wrong strand
 Seq primer: 40UP from Gibco
 High quality sequence stop: 441.
 Location/Qualifiers
 1..638
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="2800695"
 /clone_lib="NCI-CGAP-Col4"
 /tissue_type="moderately-differentiated adenocarcinoma"
 /lab_host="DH10B"
 /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.7 kb. Life Technologies catalog #:
 11531-019"

BASE COUNT 199 a 152 c 154 g 133 t
 ORIGIN

Query Match 92.0%: Score 18.4: DB 10: Length 638;
Best Local Similarity 95.0%: Pred. No. 2.5e+02;
Matches 19: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

Oy 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 150 GGATGTTGATCAGGCTCTGC 131

RESULT 9
BM943312/c 706 bp mRNA linear EST 14-MAR-2002
LOCUS UI-M-EHDP-bvn-k-16-0-UI.1 NIH_BMAP_EHDP Mus musculus cDNA clone
DEFINITION IMAGE:5694471 5', mRNA sequence.
ACCESSION BM943312 GI:19426897
VERSION
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 706)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pyx-5.
Location/Qualifiers
1..706
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5694471"
/clone_1lb="NIH_BMAP_EHDP"
/tissue_type="whole brain"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (TI phage resistant)"
/note="Organ: brain; Vector: pyx-Asc; Site: 1: EcoR I;
Site_2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pyx-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CAGCAGCAGC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP). Gene Discovery in the Developing Mouse Nervous
System, supported by National Institute of Mental Health
(NIMH), Hemin Chin, Ph.D., Program coordinator."

BASE COUNT 195 a 172 c 208 g 131 t

ORIGIN

Query Match 92.0%: Score 18.4: DB 14: Length 706;
Best Local Similarity 95.0%: Pred. No. 2.6e+02;
Matches 19: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

Oy 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 33 GGATGTTGATCAGGCTCTGC 14

RESULT 10
B1697529/c 738 bp mRNA linear EST 18-SEP-2001
LOCUS 603349086F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5376548 5',
DEFINITION mRNA sequence.
ACCESSION B1697529
VERSION B1697529.1 GI:15660158
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 738)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM1958 row: 1 column: 21
High quality sequence stop: 482.
Location/Qualifiers
1..738
/organism="Mus musculus"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:5376548"
/clone_1lb="NCI_CGAP_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith, NIH"

BASE COUNT 227 a 184 c 204 g 123 t

ORIGIN

Query Match 92.0%: Score 18.4: DB 13: Length 738;
Best Local Similarity 95.0%: Pred. No. 2.6e+02;
Matches 19: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

Oy 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 346 GGATGTTGATCAGGCTCTGC 327

RESULT 11
BG385603/c 759 bp mRNA linear EST 12-MAR-2001
LOCUS 602453904F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4582297 5',
DEFINITION mRNA sequence.
ACCESSION BG385603
VERSION BG385603.1 GI:13278469
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 759)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov

Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1304 row: 9 column: 02
High quality sequence stop: 714.

FEATURES

Location/Qualifiers

1..759

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4582297"

/clone_lib="NIH-MGC_15"

/tissue_type="adenocarcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: colon; Vector: pOTB; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dt priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 248 a 153 c 204 g 154 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 12; Length 759;

Best Local Similarity 95.0%; Pred. No. 2.7e+02;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGATGTCATCAGGCTCTGC 20

Db 248 GGATGTCATCAGGCTCTGC 229

RESULT 12

LOCUS B1688003 768 bp mRNA linear EST 18-SEP-2001

DEFINITION 603315768F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5355689 5',

mRNA sequence.

ACCESSION B1688003

VERSION B1688003.1 GI:15650631

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 768)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Jeffrey Green M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM11904 row: 9 column: 18

High quality sequence stop: 766.

FEATURES

Location/Qualifiers

1..768

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:5353689"

/clone_lib="NCI_CGAP_Mam6"

/sex="female, virgin"

/tissue_type="infiltrating ductal carcinoma"

/dev_stage="5 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies, Investigator

providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 232 a 206 c 204 g 126 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 13; Length 768;

Best Local Similarity 95.0%; Pred. No. 2.7e+02;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGATGTCATCAGGCTCTGC 20

Db 391 GGATGTCATCAGGCTCTGC 372

RESULT 13

LOCUS B1695451 778 bp mRNA linear EST 18-SEP-2001

DEFINITION 60334645F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5374261 5',

mRNA sequence.

ACCESSION B1695451

VERSION B1695451.1 GI:15658080

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 778)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM11952 row: m column: 14

High quality sequence stop: 773.

Location/Qualifiers

1..778

/organism="Mus musculus"

/strain="FVB/N-3"

/db_xref="taxon:10090"

/clone="IMAGE:5374261"

/clone_lib="NCI_CGAP_Mam2"

/tissue_type="tumor, biopsy sample"

/dev_stage="5 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies, Investigator

providing samples: Gilbert Smith, NIH"

BASE COUNT 232 a 208 c 208 g 130 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 13; Length 778;

Best Local Similarity 95.0%; Pred. No. 2.7e+02;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GGATGTCATCAGGCTCTGC 20

Db 412 GGATGTCATCAGGCTCTGC 393

RESULT 14

B1155204/c

LOCUS B1155204 824 bp mRNA linear EST 05-JUL-2001
 DEFINITION 602903188F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5032965 5',
 mRNA sequence.
 ACCESSION B1155204
 VERSION B1155204.1 GI:14615205
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 824)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LRAM11091 row: p column: 22
 High quality sequence stop: 807.
 Location/Qualifiers
 1..824
 /organism="Mus musculus"
 /strain="129,C57BL/6J,FVB/N"
 /db_xref="taxon:10090"
 /clone_image="5032965"
 /clone_id="NCI_CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt;
 Site: 2; NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
 Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999)."
 22, 37-43 (1999)."
 BASE COUNT 238 a 201 c 234 g 149 t 2 others
 ORIGIN
 Query Match 92.0%; Score 18.4; DB 13; Length 824;
 Best Local Similarity 95.0%; Pred. No. 2.8e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GGATGTTGATGAGGCTCTGC 20
 ||||||||| |||||||||
 Db 207 GGATGTTGATGAGGCTCTGC 188
 RESULT 15
 LOCUS BC028426 4072 bp mRNA linear HTC 25-APR-2002
 DEFINITION Homo sapiens, clone IMAGE:4820998, mRNA.
 ACCESSION BC028426
 VERSION BC028426.1 GI:20306365
 KEYWORDS HTC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 4072)
 Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (23-APR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 CONTACT: MGC help desk

Email: cgapbs-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shitaki
 Toshiyuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadans@systemsbio.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kerteman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAP Plate: 46 Row: h Column: 8
 This clone has the following problem: frame shifted.
 Location/Qualifiers
 1..4072
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="4820998"
 /tissue_type="Brain, hippocampus"
 /clone_id="NIH MGC_95"
 /lab_host="DH10B"
 /note="Vector: pBluescript"
 BASE COUNT 1119 a 957 c 1133 g 863 t
 ORIGIN
 Query Match 92.0%; Score 18.4; DB 11; Length 4072;
 Best Local Similarity 95.0%; Pred. No. 5.5e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GGATGTTGATGAGGCTCTGC 20
 ||||||||| |||||||||
 Db 1273 GGATGTTGATGAGGCTCTGC 1254

Search completed: November 13, 2002, 04:01:15
 Job time : 120.147 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 ; Search time 2.03468 Seconds
(without alignments)
3487.380 Million cell updates/sec

Title: US-09-697-123b-26

Perfect score: 20

Sequence: 1 ggaatgctgacagagctctgc 20

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PC7_NEM_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PC7US_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.4	92.0	477	10	US-09-563-817-166
2	18.4	92.0	348	10	US-09-917-800A-1477
3	18.4	92.0	4530	12	US-10-044-090-357
4	16.4	82.0	291	10	US-09-815-242-7923
5	15.8	79.0	76	10	US-09-864-761-21019
6	15.8	79.0	237	10	US-09-864-761-22613
7	15.8	79.0	280	10	US-09-960-352-6294
8	15.8	79.0	338	10	US-09-864-761-4266
9	15.8	79.0	434	10	US-09-960-352-4809
10	15.8	79.0	483	10	US-09-864-761-10561
11	15.8	79.0	486	10	US-09-864-761-5853
12	15.8	79.0	508	10	US-09-864-761-27182
13	15.8	79.0	1976	10	US-09-864-761-3455
14	15.4	77.0	1992	10	US-09-815-242-7772
15	15.2	76.0	1995	10	US-09-841-132-381
16	15.2	76.0	2374	10	US-09-892-985-5
17	15.2	76.0	2596	10	US-09-823-849A-408
18	15.2	76.0	2607	10	US-09-815-242-4067
19	15.2	76.0	2880	12	US-10-044-090-361

C	20	15	75.0	1534	10	US-09-350-874-7	Sequence 7, App11
C	21	15	75.0	1856	10	US-09-350-874-3	Sequence 3, App11
C	22	14.8	74.0	157	10	US-09-864-761-24969	Sequence 24969, A
C	23	14.8	74.0	296	10	US-09-294-093B-2249	Sequence 2249, Ap
C	24	14.8	74.0	383	10	US-09-960-352-11847	Sequence 11847, A
C	25	14.8	74.0	475	10	US-09-864-761-3510	Sequence 3510, Ap
C	26	14.8	74.0	551	10	US-09-604-287A-248	Sequence 248, App
C	27	14.8	74.0	551	10	US-09-339-338-248	Sequence 248, App
C	28	14.8	74.0	551	12	US-10-007-805-248	Sequence 248, App
C	29	14.8	74.0	558	10	US-09-864-761-9973	Sequence 9973, Ap
C	30	14.8	74.0	569	10	US-09-864-761-8244	Sequence 8244, Ap
C	31	14.8	74.0	583	10	US-09-864-761-8224	Sequence 8234, Ap
C	32	14.8	74.0	633	10	US-09-988-598-1559	Sequence 1569, Ap
C	33	14.8	74.0	663	10	US-09-833-381-553	Sequence 553, App
C	34	14.8	74.0	934	10	US-09-728-952-61	Sequence 61, App1
C	35	14.8	74.0	1125	9	US-09-938-842A-2333	Sequence 2333, Ap
C	36	14.8	74.0	1167	10	US-09-940-921B-6	Sequence 6, App11
C	37	14.8	74.0	1196	10	US-09-822-830A-445	Sequence 445, App
C	38	14.8	74.0	1197	10	US-09-940-921B-8	Sequence 8, App11
C	39	14.8	74.0	1200	10	US-09-815-242-4132	Sequence 4132, Ap
C	40	14.8	74.0	1744	10	US-09-940-921B-10	Sequence 10, App1
C	41	14.8	74.0	1778	10	US-09-962-832-128	Sequence 128, App
C	42	14.8	74.0	2383	10	US-09-925-302-317	Sequence 317, App
C	43	14.8	74.0	5156	10	US-09-735-705-130	Sequence 130, App
C	44	14.8	74.0	5156	10	US-09-850-716A-130	Sequence 130, App
C	45	14.8	74.0	5156	10	US-09-897-778-130	Sequence 130, App

ALIGNMENTS

RESULT 1
US-09-563-817-166/C
Sequence 166, Application US/09563817
Patent No. US20020095031A1
GENERAL INFORMATION:
APPLICANT: Nehls, Michael C.
APPLICANT: Zambrowicz, Brian
TITLE OF INVENTION: Sands, Arthur T.
TITLE OF INVENTION: Polypeptides Encoded Thereby
FILE REFERENCE: LEX-0021-USA
CURRENT APPLICATION NUMBER: US/09/563,817
PRIORITY FILING DATE: 2001-06-11
PRIOR APPLICATION NUMBER: US 60/132,343
PRIORITY FILING DATE: 1999-05-04
NUMBER OF SEQ ID NOS: 1008
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 166
LENGTH: 477
TYPE: DNA
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)...(477)
OTHER INFORMATION: n = A,T,C or G
US-09-563-817-166

Query Match 92.0%; Score 18.4; DB 10; Length 477;
Best Local Similarity 95.0%; Pred. No. 3.7;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGATGTTGATCAGGCTCTGC 20
II ||||||||||||||||
Db 214 GGGGTGTCAGCAGGCTCTGC 195

RESULT 2
US-09-917-800A-1477/C
Sequence 1477, Application US/09917800A
Patent No. US20020119462A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna


```

: APPLICANT: Porter, Mark
: APPLICANT: Johnson, Kory
: APPLICANT: Castle, Arthur
: APPLICANT: Elashoff, Michael
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Molecular Toxicology Modeling
: FILE REFERENCE: 44921-5038-US
: CURRENT APPLICATION NUMBER: US/09/917,800A
: CURRENT FILING DATE: 2001-07-31
: PRIOR APPLICATION NUMBER: US 60/222,040
: PRIOR FILING DATE: 2000-07-31
: PRIOR APPLICATION NUMBER: US 60/222,880
: PRIOR FILING DATE: 2000-11-02
: PRIOR APPLICATION NUMBER: US 60/290,029
: PRIOR FILING DATE: 2001-05-11
: PRIOR APPLICATION NUMBER: US 60/290,645
: PRIOR FILING DATE: 2001-05-15
: PRIOR APPLICATION NUMBER: US 60/292,336
: PRIOR FILING DATE: 2001-05-22
: PRIOR APPLICATION NUMBER: US 60/295,798
: PRIOR FILING DATE: 2001-06-06
: PRIOR APPLICATION NUMBER: US 60/297,457
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/298,884
: PRIOR FILING DATE: 2001-06-19
: PRIOR APPLICATION NUMBER: US 60/303,459
: PRIOR FILING DATE: 2001-07-09
: NUMBER OF SEQ ID NOS: 1740
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1477
: LENGTH: 3348
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020119462A1 U75916
: NAME/KEY: misc.feature
: LOCATION: (1)..(3339)
: OTHER INFORMATION: n = a or c or g or t
: US-09-917-800A-1477

Query Match          92.0%: Score 18.4: DB 10: Length 3348:
Best Local Similarity 95.0%: Pred. No. 4.6:
Matches 19: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

Qy      1 GCATGTTGATCAGGCTCTGC 20
        |||
Db      60 GCATGTTGATCAGGCTCTGC 41

RESULT 3
US-10-044-090-357/c
: Sequence 357, Application US/10044090
: Patent No. US20020137081A1
: GENERAL INFORMATION:
: APPLICANT: Olga Bandman
: TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
: FILE REFERENCE: PA-0028 US
: CURRENT APPLICATION NUMBER: US/10/044,090
: CURRENT FILING DATE: 2002-01-09
: NUMBER OF SEQ ID NOS: 850
: SOFTWARE: PERL Program
: SEQ ID NO 357
: LENGTH: 4530
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No. US20020137081A1 364726.12
: US-10-044-090-357

Query Match          92.0%: Score 18.4: DB 12: Length 4530:
Best Local Similarity 95.0%: Pred. No. 4.7:
Matches 19: Conservative 0: Mismatches 1: Indels 0: Gaps 0:
```

```

Qy      1 GCATGTTGATCAGGCTCTGC 20
        |||
Db      1306 GCATGTTGATCAGGCTCTGC 1287

RESULT 4
US-09-815-242-7923/c
: Sequence 7923, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes In
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7923
: LENGTH: 291
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(291)
: US-09-815-242-7923

Query Match          82.0%: Score 16.4: DB 10: Length 291:
Best Local Similarity 94.4%: Pred. No. 32:
Matches 17: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

Qy      1 GCATGTTGATCAGGCTCT 18
        |||
Db      103 GCATGTTGATCAGGCTCT 86

RESULT 5
US-09-864-761-21019
: Sequence 21019, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David A.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: FILE REFERENCE: Aecm1ca-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
```

```

: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
: SEQ ID NO 21019
: LENGTH: 76
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC012299.1
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.1
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.5
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.2
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.7
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.7
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.3
: OTHER INFORMATION: NT HIT: AF271735.1, EVALUE 4.00e-11
: OTHER INFORMATION: EST_HUMAN HIT: BE968839.1, EVALUE 5.00e-21
US-09-864-761-22613

Query Match          79.0%: Score 15.8; DB 10; Length 76;
Best Local Similarity 89.5%: Pred. No. 55;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 GCATGTTGATCAGGCTCTG 19
    ||||| ||||| |||||
Db 46 GCATGTTGCTCAGTCTG 64

RESULT 6
US-09-864-761-22613
: Sequence 22613, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
```

```

: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: FILE REFERENCE: Aecm1ca-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
: SEQ ID NO 22613
: LENGTH: 237
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC006329.2
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 9.5
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.1
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.3
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
: OTHER INFORMATION: SWISSPROT HIT: Q64726, EVALUE 6.00e-27
: OTHER INFORMATION: EST_HUMAN HIT: BF677910.1, EVALUE 1.00e-116
: OTHER INFORMATION: NT HIT: X70170.1, EVALUE 1.00e-116
US-09-864-761-22613

Query Match          79.0%: Score 15.8; DB 10; Length 237;
Best Local Similarity 89.5%: Pred. No. 62;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 2 GATGTTGATCAGGCTCTGC 20
    ||||| ||||| |||||
Db 2 GATGTTGCTCAGGCTCTCC 20

RESULT 7
US-09-960-352-6294/C
: Sequence 6294, Application US/09960352
: Patent No. US20020137139A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6294
; LENGTH: 280
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (115)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 27-LIB3058-037-Q1-K1-G11
; US-09-960-352-6294

Query Match          79.0%; Score 15.8; DB 10; Length 280;
Best Local Similarity 89.5%; Pred. No. 63;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GATGTTGATCAGGCTGTC 20
Db 70 GATGTTGATCAGGCTGTC 52

RESULT 8
; US-09-864-761-4266
; Sequence 4266, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aesomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; GENERAL INFORMATION:
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4266
; LENGTH: 398
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC012299.1
; OTHER INFORMATION: EXPRESSED IN BR474, SIGNAL - 3
; OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL - 3.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 2.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 3.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 3.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 2.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 3.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 2.7
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL - 3.3
; US-09-864-761-4266

Query Match          79.0%; Score 15.8; DB 10; Length 398;
Best Local Similarity 89.5%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGCTGTC 19
Db 363 GGATGTTGATCAGGCTGTC 381

RESULT 9
; US-09-960-352-4809/C
; Sequence 4809, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4809
; LENGTH: 434
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 21-LIB3057-006-Q1-K1-F9
; US-09-960-352-4809

Query Match          79.0%; Score 15.8; DB 10; Length 434;
Best Local Similarity 89.5%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GATGTTGATCAGGCTGTC 20
Db 69 GATGTTGATCAGGCTGTC 51

RESULT 10
; US-09-864-761-10561
; Sequence 10561, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10561
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121758.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.91
US-09-864-761-10561

Query Match          79.0%; Score 15.8; DB 10; Length 483;
Best Local Similarity 89.5%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCATGTTGATCAGGCTCTG 19
Db 370 GGTGTGTCAGGAGCTG 388
```

RESULT 11

```

US-09-864-761-5853
; Sequence 5853, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
```

```

; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5853
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006329.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 9.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 73
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
US-09-864-761-5853

Query Match          79.0%; Score 15.8; DB 10; Length 486;
Best Local Similarity 89.5%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GATGTTGATCAGGCTCTGC 20
Db 261 GATGTTGTCAGGCTCTCC 279
```

RESULT 12

```

US-09-864-761-27182
; Sequence 27182, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
```

```

; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/226,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27182
; LENGTH: 508
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121758.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.91
; OTHER INFORMATION: EST HUMAN HIT: AU144395.1, EVALUATE 1.20e-01
; OTHER INFORMATION: NT HIT: AJ272265.1, EVALUATE 5.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P21997, EVALUATE 1.10e+00
US-09-864-761-27182

Query Match          79.0%; Score 15.8; DB 10; Length 508;
Best Local Similarity 89.5%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCATGTTGATCAGGCTG 19
DB 400 GGTGTTGATCAGGACTG 418
```

RESULT 13

```
US-09-864-761-3455/c
; Sequence 3455, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
```

```

; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3455
; LENGTH: 1976
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006116.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.94
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95
US-09-864-761-3455

Query Match          79.0%; Score 15.8; DB 10; Length 1976;
Best Local Similarity 89.5%; Pred. No. 77;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GATGTTGATCAGGCTG 20
DB 1153 GATGTTGATCAGGTATGC 1135
```

RESULT 14

```
US-09-815-242-7772/c
; Sequence 7772, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
```

APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes In
FILE REFERENCE: ELITRA-011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7772
LENGTH: 1992
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1992)
US-09-815-242-7772

Query Match 77.0%: Score 15.4; DB 10; Length 1992;
Best Local Similarity 94.1%: Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 GGATGTCATCAGGCTC 17
||||| |||||||
DB 181 GGATGTCATCAGGCTC 165

RESULT 15
US-09-841-132-381
Sequence 381, Application US/09841132
Patent No. US20020061848A1
GENERAL INFORMATION:
APPLICANT: Bhatia, Ajay
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Probst, Peter
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
FILE REFERENCE: 210121.469C8
CURRENT APPLICATION NUMBER: US/09/841,132
CURRENT FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 599
SOFTWARE: FastSeq for Windows Version 3.0/4.0
SEQ ID NO 381
LENGTH: 1995
TYPE: DNA
ORGANISM: Chlamydia pneumoniae
US-09-841-132-381

Query Match 76.0%: Score 15.2; DB 10; Length 1995;
Best Local Similarity 85.0%: Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 GGATGTCATCAGGCTGC 20
||||| ||||||| |||||

DB 393 GGATGTCATCAGCCTCTGC 412

Search completed: November 12, 2002, 16:59:33
Job time : 5.03468 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:56:36 ; Search time 2.68208 Seconds
(without alignments)
2286.857 Million cell updates/sec

Title: US-09-697-123b-26

Perfect score: 20

Sequence: 1 ggatgttcacagggtctgc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents_NA:*
- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
 - 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
 - 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
 - 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
 - 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
 - 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
C 1	20	100.0	228	4	US-08-750-088A-69	Sequence 69, Appl
C 2	20	100.0	254	4	US-08-750-088A-37	Sequence 37, Appl
C 3	20	100.0	306	4	US-09-147-935A-1	Sequence 1, Appl
C 4	20	100.0	306	4	US-09-147-935A-3	Sequence 3, Appl
C 5	20	100.0	306	4	US-09-147-935A-4	Sequence 4, Appl
C 6	20	100.0	306	4	US-09-147-935A-5	Sequence 5, Appl
C 7	20	100.0	306	4	US-09-147-935A-8	Sequence 8, Appl
C 8	20	100.0	306	4	US-09-147-935A-9	Sequence 9, Appl
C 9	20	100.0	306	4	US-09-147-935A-10	Sequence 10, Appl
C 10	20	100.0	306	4	US-09-147-935A-11	Sequence 11, Appl
C 11	20	100.0	306	4	US-09-147-935A-12	Sequence 12, Appl
C 12	20	100.0	306	4	US-09-147-935A-13	Sequence 13, Appl
C 13	20	100.0	306	4	US-09-147-935A-14	Sequence 14, Appl
C 14	20	100.0	306	4	US-09-147-935A-15	Sequence 15, Appl
C 15	20	100.0	306	4	US-09-147-935A-17	Sequence 17, Appl
C 16	20	100.0	306	4	US-09-147-935A-18	Sequence 18, Appl
C 17	20	100.0	306	4	US-09-147-935A-22	Sequence 22, Appl
C 18	20	100.0	306	4	US-09-147-935A-27	Sequence 27, Appl
C 19	20	100.0	306	4	US-09-147-935A-28	Sequence 28, Appl
C 20	20	100.0	306	4	US-09-147-935A-29	Sequence 29, Appl
C 21	20	100.0	306	4	US-09-147-935A-30	Sequence 30, Appl
C 22	20	100.0	306	4	US-09-147-935A-31	Sequence 31, Appl
C 23	20	100.0	306	4	US-09-147-935A-32	Sequence 32, Appl
C 24	20	100.0	306	4	US-09-147-935A-33	Sequence 33, Appl
C 25	20	100.0	306	4	US-09-147-935A-35	Sequence 35, Appl
C 26	20	100.0	306	4	US-09-147-935A-36	Sequence 36, Appl
C 27	20	100.0	306	4	US-09-147-935A-37	Sequence 37, Appl

C 28	20	100.0	306	4	US-09-147-935A-38	Sequence 38, Appl
C 29	20	100.0	306	4	US-09-147-935A-39	Sequence 39, Appl
C 30	20	100.0	306	4	US-09-147-935A-40	Sequence 40, Appl
C 31	20	100.0	306	4	US-09-147-935A-43	Sequence 43, Appl
C 32	20	100.0	306	4	US-09-147-935A-46	Sequence 46, Appl
C 33	20	100.0	306	4	US-09-147-935A-47	Sequence 47, Appl
C 34	20	100.0	319	4	US-08-750-088A-35	Sequence 35, Appl
C 35	20	100.0	324	4	US-08-750-088A-36	Sequence 36, Appl
C 36	18.4	92.0	306	4	US-09-147-935A-20	Sequence 20, Appl
C 37	18.4	92.0	306	4	US-09-147-935A-21	Sequence 21, Appl
C 38	18.4	92.0	306	4	US-09-147-935A-23	Sequence 23, Appl
C 39	18.4	92.0	306	4	US-09-147-935A-25	Sequence 25, Appl
C 40	18.4	92.0	306	4	US-09-147-935A-26	Sequence 26, Appl
C 41	18.4	92.0	306	4	US-09-147-935A-34	Sequence 34, Appl
C 42	18.4	92.0	306	4	US-09-147-935A-42	Sequence 42, Appl
C 43	18.4	92.0	306	4	US-09-147-935A-44	Sequence 44, Appl
C 44	16.8	84.0	306	4	US-09-147-935A-2	Sequence 2, Appl
C 45	16.8	84.0	306	4	US-09-147-935A-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-750-088A-69/c
; Sequence 69, Application US/08750088A
; Patent No. 6329138
; GENERAL INFORMATION:
; APPLICANT: DE BEENHOUWER, HANS
; APPLICANT: PORTAELS, FRAN OISE
; APPLICANT: MACHTELINCKX, LIEVE
; APPLICANT: JANNES, GEERT
; APPLICANT: ROSSAU, RUDI
; TITLE OF INVENTION: METHOD FOR DETECTION OF THE ANTIBIOTIC
; TITLE OF INVENTION: RESISTANCE SPECTRUM OF MYCOBACTERIUM SPECIES
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,088A
; FILING DATE: 21-FEB-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDSTEIN, JORGE A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1657.0010000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 228 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-750-088A-69

Query Match 100.0%; Score 20; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATGTTGATCAGGTCCTGC 20

```
Db      88 GGATGTTGATCAGGCTGTCG 69
|||||
RESULT 2
US-08-750-088A-37/c
; Sequence 37, Application US/08750088A
; Patent No. 6329138
; GENERAL INFORMATION:
; APPLICANT: DE BEENHOUWER, HANS
; APPLICANT: PORTAELS, FRAN OISE
; APPLICANT: MACHTELINCKX, LIEVE
; APPLICANT: JANNES, GEERT
; APPLICANT: ROSSAU, GUIDI
; TITLE OF INVENTION: METHOD FOR DETECTION OF THE ANTIBIOTIC
; TITLE OF INVENTION: RESISTANCE SPECTRUM OF MYCOBACTERIUM SPECIES
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,088A
; FILING DATE: 21-FEB-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDSTEIN, JORGE A.
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1657.0010000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-750-088A-37

Query Match      100.0%; Score 20; DB 4; Length 254;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGATGTTGATCAGGCTGTCG 20
|||||
Db      94 GGATGTTGATCAGGCTGTCG 75
|||||
RESULT 3
US-09-147-935A-1/c
; Sequence 1, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/OF425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; CURRENT FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 4
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium aurum
US-09-147-935A-4

Query Match      100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGATGTTGATCAGGCTGTCG 20
|||||
Db      129 GGATGTTGATCAGGCTGTCG 110
|||||
RESULT 4
US-09-147-935A-3/c
; Sequence 3, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/OF425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; CURRENT FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 3
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium asiaticum
US-09-147-935A-3

Query Match      100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GGATGTTGATCAGGCTGTCG 20
|||||
Db      129 GGATGTTGATCAGGCTGTCG 110
|||||
RESULT 5
US-09-147-935A-4/c
; Sequence 4, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/OF425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; CURRENT FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 4
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium aurum
US-09-147-935A-4

Query Match      100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
Oy 1 GGATGTTGATCAGGGTCTGC 20
|||||
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 6
US-09-147-935A-5/c
; Sequence 5, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; FILE REFERENCE: 0136/OF425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; CURRENT FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 5
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium avium
; US-09-147-935A-5

Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGATGTTGATCAGGGTCTGC 20
|||||
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 7
US-09-147-935A-8/c
; Sequence 8, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; FILE REFERENCE: 0136/OF425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; CURRENT FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 8
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium celatum Typel
; US-09-147-935A-8

Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGATGTTGATCAGGGTCTGC 20
|||||
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 8
US-09-147-935A-9/c
; Sequence 9, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
```

```
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; FILE REFERENCE: 0136/OF425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; CURRENT FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 9
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium celatum Type2
; US-09-147-935A-9
```

```
Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 GGATGTTGATCAGGGTCTGC 20
|||||
Db 129 GGATGTTGATCAGGGTCTGC 110
```

```
RESULT 9
US-09-147-935A-10/c
; Sequence 10, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; FILE REFERENCE: 0136/OF425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; CURRENT FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 10
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium chelonae
; US-09-147-935A-10
```

```
Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 GGATGTTGATCAGGGTCTGC 20
|||||
Db 129 GGATGTTGATCAGGGTCTGC 110
```

```
RESULT 10
US-09-147-935A-11/c
; Sequence 11, Application US/09147935A
; Patent No. 6242584
; GENERAL INFORMATION:
; APPLICANT: KOOK, Yoon-Hoh
; APPLICANT: KIM, Bum-Joon
; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY
; FILE REFERENCE: 0136/OF425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; CURRENT FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/KR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 11
```

; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium chitae
US-09-147-935A-11

Query Match 100.0%; Score 20; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20
|||||

Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 11

US-09-147-935A-12/c

; Sequence 12, Application US/09147935A

; Patent No. 6242584

; GENERAL INFORMATION:

; APPLICANT: KOOK, Yoon-Hoh

; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY

; FILE REFERENCE: 0136/0F425

; CURRENT APPLICATION NUMBER: US/09/147,935A

; PRIOR FILING DATE: 1999-03-19

; PRIOR FILING DATE: 1998-07-28

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: KOPATIN 1.0

; SEQ ID NO 12

; LENGTH: 306

; TYPE: DNA

; ORGANISM: Mycobacterium fallax

US-09-147-935A-12

Query Match

100.0%; Score 20; DB 4; Length 306;

Best Local Similarity 100.0%; Pred. No. 1.1;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20
|||||

Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 12

US-09-147-935A-13/c

; Sequence 13, Application US/09147935A

; Patent No. 6242584

; GENERAL INFORMATION:

; APPLICANT: KOOK, Yoon-Hoh

; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY

; FILE REFERENCE: 0136/0F425

; CURRENT APPLICATION NUMBER: US/09/147,935A

; PRIOR FILING DATE: 1999-03-19

; PRIOR FILING DATE: 1998-07-28

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: KOPATIN 1.0

; SEQ ID NO 13

; LENGTH: 306

; TYPE: DNA

; ORGANISM: Mycobacterium flavescens

US-09-147-935A-13

Query Match

100.0%; Score 20; DB 4; Length 306;

Best Local Similarity 100.0%; Pred. No. 1.1;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20
|||||

Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 13

US-09-147-935A-14/c

; Sequence 14, Application US/09147935A

; Patent No. 6242584

; GENERAL INFORMATION:

; APPLICANT: KOOK, Yoon-Hoh

; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY

; FILE REFERENCE: 0136/0F425

; CURRENT APPLICATION NUMBER: US/09/147,935A

; PRIOR FILING DATE: 1999-03-19

; PRIOR FILING DATE: 1998-07-28

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: KOPATIN 1.0

; SEQ ID NO 14

; LENGTH: 306

; TYPE: DNA

; ORGANISM: Mycobacterium fortuitum

US-09-147-935A-14

Query Match

100.0%; Score 20; DB 4; Length 306;

Best Local Similarity 100.0%; Pred. No. 1.1;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20
|||||

Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 14

US-09-147-935A-15/c

; Sequence 15, Application US/09147935A

; Patent No. 6242584

; GENERAL INFORMATION:

; APPLICANT: KOOK, Yoon-Hoh

; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY

; FILE REFERENCE: 0136/0F425

; CURRENT APPLICATION NUMBER: US/09/147,935A

; PRIOR FILING DATE: 1999-03-19

; PRIOR FILING DATE: 1998-07-28

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: KOPATIN 1.0

; SEQ ID NO 15

; LENGTH: 306

; TYPE: DNA

; ORGANISM: Mycobacterium fortuitum 49403

US-09-147-935A-15

Query Match

100.0%; Score 20; DB 4; Length 306;

Best Local Similarity 100.0%; Pred. No. 1.1;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20
|||||

Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 15

US-09-147-935A-17/c

; Sequence 17, Application US/09147935A

; Patent No. 6242584

; GENERAL INFORMATION:

; APPLICANT: KOOK, Yoon-Hoh

; TITLE OF INVENTION: A METHOD FOR IDENTIFYING MYCOBACTERIAL SPECIES BY

; TITLE OF INVENTION: COMPARATIVE SEQUENCE ANALYSIS OF rpoB GENE
; FILE REFERENCE: 0136/0F425
; CURRENT APPLICATION NUMBER: US/09/147,935A
; CURRENT FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: PCT/RR98/00228
; PRIOR FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: KOPATIN 1.0
; SEQ ID NO 17
; LENGTH: 306
; TYPE: DNA
; ORGANISM: Mycobacterium genavense
US-09-147-935A-17

Query Match	100.0%	Score 20;	DB 4;	Length 306;
Best Local Similarity	100.0%	Pred. No. 1.1;		
Matches	20;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGCTCTGC 20
Db 129 GGATGTTGATCAGGCTCTGC 110
|||||

Search completed: November 12, 2002, 21:03:47
Job time : 3.68208 secs

GenCore version 5.1.3

Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 : Search time 13.0983 Seconds
(without alignments)
3438.621 Million cell updates/sec

Title: US-09-697-123B-26

Perfect score: 20

Sequence: 1 ggaatgtgatcagggtctgc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :		N_Geneseq_101002.*	
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*		
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*		
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*		
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*		
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*		
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*		
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*		
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*		
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*		
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*		
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*		
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*		
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*		
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*		
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*		
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*		
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*		
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*		
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*		
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*		
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*		
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*		
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*		
24:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	20	100.0	20 22 AAS05226
C 2	20	100.0	228 17 AAT29617
C 3	20	100.0	254 17 AAT29615
C 4	20	100.0	300 24 AAS99529
C 5	20	100.0	306 19 AAX27208
C 6	20	100.0	306 19 AAX27209
C 7	20	100.0	306 19 AAX27210
C 8	20	100.0	306 19 AAX27211
C 9	20	100.0	306 19 AAX27212

1	20	100.0	20 22 AAS05226	Mycobacterium spec
C 2	20	100.0	228 17 AAT29617	Partial sequence o
C 3	20	100.0	254 17 AAT29615	Partial sequence o
C 4	20	100.0	300 24 AAS99529	Mycobacterium spec
C 5	20	100.0	306 19 AAX27208	RpoB gene fragment
C 6	20	100.0	306 19 AAX27209	RpoB gene fragment
C 7	20	100.0	306 19 AAX27210	RpoB gene fragment
C 8	20	100.0	306 19 AAX27211	RpoB gene fragment
C 9	20	100.0	306 19 AAX27212	RpoB gene fragment

XX PS Disclosure; Page 13; 50pp; English.

XX CC The present sequence for Mycobacterium species rpoB gene fragment PCR

XX CC primer RPO3' is used with PCR primer RPO5' (AAS05225) to amplify the

XX CC rpoB gene region from 24 rpoB gene fragments (AAS05201-AAS05224) from

XX CC various Mycobacterial species. These rpoB gene fragments can be used

XX CC in the diagnosis and identification of Mycobacterium species using a

XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)

XX CC method. The method comprises obtaining a restriction fragment length

XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,

XX CC amplifying and digesting the DNA fragment from the microorganism to

XX CC be identified and comparing the RFLP patterns from the known rpoB gene

XX CC fragments with the unidentified fragment. The rpoB gene fragments

XX CC are useful to identify a wide range of Mycobacterium species, e.g. for

XX CC diagnosis or to obtain epidemiological and pathogenesis information for

XX CC selection of appropriate therapies including M. tuberculosis, M. leprae

XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected

XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene

XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR

XX CC required), and can differentiate between many species in a single

XX CC experiment, including those difficult to distinguish by usual biochemical

XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for

XX CC detecting specific Mycobacterial species.

SQ Sequence 20 BP; 3 A; 3 C; 8 G; 6 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 20;

Best Local Similarity 100.0%; Pred. No. 2.3; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0;

QY 1 GGATGTTGATCAGGGTCTGC 20

|||||

Db 1 GGATGTTGATCAGGGTCTGC 20

RESULT 2

AAT29617/c

ID AAT29617 standard; DNA; 228 BP.

XX AC AAT29617;

XX DT 11-JUL-1996 (first entry)

XX DE Partial sequence of Mycobacterium MAC strain ITG 926 rpoB gene.

XX KW Antibiotic; resistance; spectrum; gene; mycobacterium;

XX KW determination; amplification; MAC strain; rpoB; fragment;

XX KW probe; differential; hybridisation; pattern; rifampicin;

XX KW rifabutin; species identification; ss.

XX OS Mycobacterium MAC strain.

XX PN WO9533851-A2.

XX PD 14-DEC-1995.

XX PF 09-JUN-1995; 95WO-EP02230.

XX PR 09-JUN-1994; 94EP-0870093.

XX PA (INNO-) INNOGENETICS NV.

XX PI De Beenhouwer H, Jannes G, Machtelinckx L, Portael F;

XX PI Rossau R;

XX DR WPI; 1996-040250/04.

XX PT Probes and primers for determin. of antibiotic resistance spectrum of

XX PT Mycobacterium, opt. coupled with species identification - from

XX PT different patterns of hybridisation with rpoB gene

XX PS Claim 13; Fig 11; 69pp; English.

XX

CC The antibiotic resistance spectrum (ARS) of a mycobacterium can be

CC determined by amplifying the relevant part of the antibiotic

CC resistance gene, hybridising it with at least 1 rpoB gene probe,

CC detecting the hybrids formed and inferring the ARS, and opt. the

CC spp. using species specific probes (i.e. AAT12148/49 derived from

CC the present sequence, the partial nucleotide sequence of the

CC presumptive Mycobacterium MAC strain ITG 926 rpoB gene) from the

CC differential hybridisation patterns. The method is partic. useful

CC for the detection of rifampicin and/or rifabutin resistance in

CC M. leprae or M. tuberculosis, and mycobacterial spp.

CC identification. The method is rapid and reliable and provides

CC simultaneous determin. of ARS and spp. identity.

XX SQ Sequence 228 BP; 40 A; 76 C; 73 G; 37 T; 2 other;

Query Match 100.0%; Score 20; DB 17; Length 228;

Best Local Similarity 100.0%; Pred. No. 3.1; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0;

QY 1 GGATGTTGATCAGGGTCTGC 20

|||||

Db 88 GGATGTTGATCAGGGTCTGC 69

RESULT 3

AAT29615/c

ID AAT29615 standard; DNA; 254 BP.

XX AC AAT29615;

XX DT 11-JUL-1996 (first entry)

XX DE Partial sequence of M. scrofulaceum strain ITG 4979 rpoB gene.

XX KW Antibiotic; resistance; spectrum; gene; mycobacterium;

XX KW determination; amplification; scrofulaceum; rpoB; fragment;

XX KW probe; differential; hybridisation; pattern; rifampicin;

XX KW rifabutin; species identification; ss.

XX OS Mycobacterium scrofulaceum.

XX PN WO9533851-A2.

XX PD 14-DEC-1995.

XX PF 09-JUN-1995; 95WO-EP02230.

XX PR 09-JUN-1994; 94EP-0870093.

XX PA (INNO-) INNOGENETICS NV.

XX PI De Beenhouwer H, Jannes G, Machtelinckx L, Portael F;

XX PI Rossau R;

XX DR WPI; 1996-040250/04.

XX PT Probes and primers for determin. of antibiotic resistance spectrum of

XX PT Mycobacterium, opt. coupled with species identification - from

XX PT different patterns of hybridisation with rpoB gene

XX PS Claim 13; Fig 7; 69pp; English.

XX CC The antibiotic resistance spectrum (ARS) of a mycobacterium can be

XX CC determined by amplifying the relevant part of the antibiotic

XX CC resistance gene, hybridising it with at least 1 rpoB gene probe,

XX CC detecting the hybrids formed and inferring the ARS, and opt. the

XX CC spp. using species specific probes (i.e. AAT12146 derived from the

XX CC present sequence, the partial nucleotide sequence of the

XX CC presumptive M. scrofulaceum strain ITG 4979 rpoB gene) from the

XX CC differential hybridisation patterns. The method is partic. useful

XX CC for the detection of rifampicin and/or rifabutin resistance in

XX CC M. leprae or M. tuberculosis, and mycobacterial spp.

CC Identification. The method is rapid and reliable and provides
CC simultaneous determin. of ARS and spp. identity.
XX

SQ Sequence 254 BP; 40 A; 84 C; 90 G; 39 T; 1 other;

Query Match 100.0%; Score 20; DB 17; Length 254;
Best Local Similarity 100.0%; Pred. No. 3.1; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0;

OY 1 GGATGTTGATCAGGGTCTGC 20
|||||
DB 94 GGATGTTGATCAGGGTCTGC 75

RESULT 4

AAS99529/c
ID AAS99529 standard; DNA; 300 BP.

XX AC AAS99529;

XX DT 12-MAR-2002 (first entry)

DE Mycobacterium species identification primer #4.

XX Drug resistance detection; mycobacterial species identification; probe;
KW oligonucleotide chip; rpoB; sputum; blood; cerebrospinal fluid; ss,
KW primer.

OS Mycobacterium avium.

PN WO200192573-A1.

XX PD 06-DEC-2001.

XX PF 30-MAY-2001; 2001WO-KR00904.

XX PR 30-MAY-2000; 2000KR-0029369.

XX PA (BIOM-) BIOMEDLAB CO LTD.

PI Kim H, Kim N, Yoon S, Kim J, Park M;

XX WPI; 2002-075472/10.

XX Kit for mycobacterial species identification and drug resistance
PT detection, has oligonucleotide chip with species identification probe,
PT a mycobacterial drug-resistance detection probe, and its contrast group
PT probe.

PS Disclosure; Page 21; 74pp; English.

XX The invention relates to a diagnostic kit for mycobacterial species
CC identification and drug resistance detection comprising an
CC oligonucleotide chip including a species identification probe, a
CC mycobacterial drug-resistance detection probe, a contrast group probe
CC corresponding to each drug resistance detection probe, and a marker for
CC detecting a hybridisation of the oligonucleotide chip and a specimen. The
CC identification probe is comprised of species-specific DNA sequences of
CC mycobacterial rpoB gene and the detection probe is comprised of one or
CC more modified codons of mycobacterial rpoB gene. The method involves
CC amplifying rpoB gene fragments of specimen by Polymerase Chain Reaction
CC (PCR) and discriminating species by fluorescent intensity corresponding
CC to a particular species. The specimen is preferably uncultured sputum,
CC blood or cerebrospinal fluid of a patient. Sequences AAS99478-AAS99569
CC represent mycobacterium species identification probes and primers of the
CC invention.

SQ Sequence 300 BP; 49 A; 105 C; 101 G; 45 T; 0 other;

Query Match 100.0%; Score 20; DB 24; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGATGTTGATCAGGGTCTGC 20
|||||
DB 129 GGATGTTGATCAGGGTCTGC 110

RESULT 5

AAX27208/c
ID AAX27208 standard; DNA; 306 BP.

XX AC AAX27208;

XX DT 27-MAY-1999 (first entry)

XX DE RpoB gene fragment.

XX RpoB gene; mycobacteria; phylogenetic tree construction;

KW mycobacterial species identification; phylogenetic analysis; ss.

XX OS Mycobacteria simiae.

XX PN WO9905316-A1.

XX PD 04-FEB-1999.

XX PF 28-JUL-1998; 98KR-0000228.

XX PR 28-JUL-1997; 97KR-0035501.

XX PA (BION-) BIONEER CORP.

XX PI Kim B, Kook Y;

XX WPI; 1998-539367/46.

XX New pair of polymerase chain reaction (PCR) primers - for
PT sequence-specific amplification of the rpoB gene from mycobacterial
PT species, useful for detecting and identifying mycobacterial species
XX
XX Claim 37; Page 73-74; 91pp; English.

XX This sequence represents a mycobacterial rpoB gene fragment, that is
CC amplified using the PCR primers of the invention. The primers form a
CC method of detecting and identifying mycobacterial species by constructing
CC a phylogenetic tree for the species. The use of the primers for
CC sequence-specific amplification of the rpoB gene (encoding the beta
CC subunit of RNA polymerase) from mycobacterial species provides an
CC efficient way of characterising these species. In addition to
CC phylogenetic analysis, the rpoB gene can be used as an alternative to
CC the 16S rRNA gene because it has four subunits, which are highly
CC conserved throughout prokaryotes. The method is particularly useful for
CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
CC susceptibility can be simultaneously determined in M. tuberculosis.

SQ Sequence 306 BP; 53 A; 101 C; 100 G; 52 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGATGTTGATCAGGGTCTGC 20
|||||
DB 129 GGATGTTGATCAGGGTCTGC 110

RESULT 6

AAX27209/c
ID AAX27209 standard; DNA; 306 BP.

XX AC AAX27209;

XX DT 27-MAY-1999 (first entry)

XX DE RpoB gene fragment.

```

XX RpoB gene; mycobacteria; phylogenetic tree construction;
KW mycobacterial species identification; phylogenetic analysis; ss.
XX
OS Mycobacteria smegmatis.
XX
PN WO9905316-A1.
XX
PD 04-FEB-1999.
XX
XX 28-JUL-1998; 98KR-0000228.
PF
XX
PR 28-JUL-1997; 97KR-0035501.
XX
XX (BION-) BIONEER CORP.
PA
XX Kim B, Kook Y;
PI
XX WPI; 1998-539367/46.
DR
XX
XX New pair of polymerase chain reaction (PCR) primers - for
PT sequence-specific amplification of the rpoB gene from mycobacterial
PT species, useful for detecting and identifying mycobacterial species
XX
PS Claim 38; Page 74; 91pp; English.
XX
CC This sequence represents a mycobacterial rpoB gene fragment, that is
CC amplified using the PCR primers of the invention. The primers form a
CC method of detecting and identifying mycobacterial species by constructing
CC a phylogenetic tree for the species. The use of the primers for
CC sequence-specific amplification of the rpoB gene (encoding the beta
CC subunit of RNA polymerase) from mycobacterial species provides an
CC efficient way of characterising these species. In addition to
CC phylogenetic analysis, the rpoB gene can be used as an alternative to
CC the 16S rRNA gene because it has four subunits, which are highly
CC conserved throughout prokaryotes. The method is particularly useful for
CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
CC susceptibility can be simultaneously determined in M. tuberculosis.
XX
SQ Sequence 306 BP; 50 A; 108 C; 96 G; 52 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGCTCTGC 20
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 7
AAX27210/C
ID AAX27210 standard; DNA; 306 BP.
XX
XX AAX27210;
AC
XX
XX 27-MAY-1999 (first entry)
DT
XX
XX RpoB gene fragment.
DE
XX
XX RpoB gene; mycobacteria; phylogenetic tree construction;
KW mycobacterial species identification; phylogenetic analysis; ss.
XX
XX Mycobacteria szulgai.
OS
XX
XX WO9905316-A1.
PN
XX
XX 04-FEB-1999.
PD
XX
XX 28-JUL-1998; 98KR-0000228.
PF
XX
XX 28-JUL-1997; 97KR-0035501.
PR
XX

```

```

PA (BION-) BIONEER CORP.
XX
XX Kim B, Kook Y;
PI
XX WPI; 1998-539367/46.
DR
XX
XX New pair of polymerase chain reaction (PCR) primers - for
PT sequence-specific amplification of the rpoB gene from mycobacterial
PT species, useful for detecting and identifying mycobacterial species
XX
PS Claim 39; Page 74; 91pp; English.
XX
XX This sequence represents a mycobacterial rpoB gene fragment, that is
CC amplified using the PCR primers of the invention. The primers form a
CC method of detecting and identifying mycobacterial species by constructing
CC a phylogenetic tree for the species. The use of the primers for
CC sequence-specific amplification of the rpoB gene (encoding the beta
CC subunit of RNA polymerase) from mycobacterial species provides an
CC efficient way of characterising these species. In addition to
CC phylogenetic analysis, the rpoB gene can be used as an alternative to
CC the 16S rRNA gene because it has four subunits, which are highly
CC conserved throughout prokaryotes. The method is particularly useful for
CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
CC susceptibility can be simultaneously determined in M. tuberculosis.
XX
SQ Sequence 306 BP; 50 A; 104 C; 102 G; 50 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGCTCTGC 20
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 8
AAX27211/C
ID AAX27211 standard; DNA; 306 BP.
XX
XX AAX27211;
AC
XX
XX 27-MAY-1999 (first entry)
DT
XX
XX RpoB gene fragment.
DE
XX
XX RpoB gene; mycobacteria; phylogenetic tree construction;
KW mycobacterial species identification; phylogenetic analysis; ss.
XX
XX Mycobacteria terrae.
OS
XX
XX WO9905316-A1.
PN
XX
XX 04-FEB-1999.
PD
XX
XX 28-JUL-1998; 98KR-0000228.
PF
XX
XX 28-JUL-1997; 97KR-0035501.
PR
XX
XX (BION-) BIONEER CORP.
PA
XX
XX Kim B, Kook Y;
PI
XX
XX WPI; 1998-539367/46.
DR
XX
XX New pair of polymerase chain reaction (PCR) primers - for
PT sequence-specific amplification of the rpoB gene from mycobacterial
PT species, useful for detecting and identifying mycobacterial species
XX
PS Claim 40; Page 74-75; 91pp; English.
XX
XX This sequence represents a mycobacterial rpoB gene fragment, that is
CC amplified using the PCR primers of the invention. The primers form a

```

CC method of detecting and identifying mycobacterial species by constructing
 CC a phylogenetic tree for the species. The use of the primers for
 CC sequence-specific amplification of the rpoB gene (encoding the beta
 CC subunit of RNA polymerase) from mycobacterial species provides an
 CC efficient way of characterising these species. In addition to
 CC the 16S rRNA gene, the rpoB gene can be used as an alternative to
 CC phylogenetic analysis, the rpoB gene has four subunits, which are highly
 CC conserved throughout prokaryotes. The method is particularly useful for
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
 CC susceptibility can be simultaneously determined in *M. tuberculosis*.
 CC
 XX Sequence 306 BP; 49 A; 102 C; 103 G; 52 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;
 Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGGTCTGC 20
 Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 9
 AAX27212/c
 ID AAX27212 standard; DNA; 306 BP.
 AC AAX27212;
 XX
 XX 27-MAY-1999 (first entry)
 DT RpoB gene fragment.
 DE RpoB gene: mycobacteria; phylogenetic tree construction;
 XX mycobacterial species identification; phylogenetic analysis; ss.
 KW Mycobacteria thermoresistable.
 OS
 XX WO9905316-A1.
 PN 04-FEB-1999.
 PD 28-JUL-1998; 98KR-0000228.
 PF 28-JUL-1997; 97KR-0035501.
 XX (BION-) BIONEER CORP.
 PA Kim B, Kook Y;
 XX WPI; 1998-539367/46.

PT New pair of polymerase chain reaction (PCR) primers - for
 PT sequence-specific amplification of the rpoB gene from mycobacterial
 PT species, useful for detecting and identifying mycobacterial species
 XX
 PS Claim 41; Page 75; 91pp; English.

CC This sequence represents a mycobacterial rpoB gene fragment, that is
 CC amplified using the PCR primers of the invention. The primers form a
 CC method of detecting and identifying mycobacterial species by constructing
 CC a phylogenetic tree for the species. The use of the primers for
 CC sequence-specific amplification of the rpoB gene (encoding the beta
 CC subunit of RNA polymerase) from mycobacterial species provides an
 CC efficient way of characterising these species. In addition to
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to
 CC the 16S rRNA gene because it has four subunits, which are highly
 CC conserved throughout prokaryotes. The method is particularly useful for
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
 CC susceptibility can be simultaneously determined in *M. tuberculosis*.
 CC
 XX Sequence 306 BP; 51 A; 111 C; 102 G; 42 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;

Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGGTCTGC 20
 Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 10
 AAX27213/c
 ID AAX27213 standard; DNA; 306 BP.
 XX
 AC AAX27213;
 XX
 XX 27-MAY-1999 (first entry)
 DT RpoB gene fragment.
 DE RpoB gene: mycobacteria; phylogenetic tree construction;
 KW mycobacterial species identification; phylogenetic analysis; ss.
 OS Mycobacteria triviale.
 XX
 PN WO9905316-A1.
 XX 04-FEB-1999.
 PD 28-JUL-1998; 98KR-0000228.
 PF 28-JUL-1997; 97KR-0035501.
 XX (BION-) BIONEER CORP.
 PA Kim B, Kook Y;
 XX WPI; 1998-539367/46.

PT New pair of polymerase chain reaction (PCR) primers - for
 PT sequence-specific amplification of the rpoB gene from mycobacterial
 PT species, useful for detecting and identifying mycobacterial species
 XX
 PS Claim 42; Page 75; 91pp; English.

CC This sequence represents a mycobacterial rpoB gene fragment, that is
 CC amplified using the PCR primers of the invention. The primers form a
 CC method of detecting and identifying mycobacterial species by constructing
 CC a phylogenetic tree for the species. The use of the primers for
 CC sequence-specific amplification of the rpoB gene (encoding the beta
 CC subunit of RNA polymerase) from mycobacterial species provides an
 CC efficient way of characterising these species. In addition to
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to
 CC the 16S rRNA gene because it has four subunits, which are highly
 CC conserved throughout prokaryotes. The method is particularly useful for
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
 CC susceptibility can be simultaneously determined in *M. tuberculosis*.
 CC
 XX Sequence 306 BP; 50 A; 104 C; 109 G; 43 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;
 Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGGTCTGC 20
 Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 11
 AAX27216/c
 ID AAX27216 standard; DNA; 306 BP.
 XX
 AC AAX27216;
 XX


```

DT 27-MAY-1999 (first entry)
DE RpoB gene fragment.
XX RpoB gene; mycobacteria; phylogenetic tree construction;
KW mycobacterial species identification; phylogenetic analysis; ss.
XX Mycobacteria vaccae.
XX WO9905316-A1.
XX 04-FEB-1999.
XX 28-JUL-1998; 98KR-0000228.
XX 28-JUL-1997; 97KR-0035501.
XX (BION-) BIONEER CORP.
XX Kim B, Kook Y;
XX WPI; 1998-539367/46.
XX New pair of polymerase chain reaction (PCR) primers - for
PT sequence-specific amplification of the rpoB gene from mycobacterial
PT species, useful for detecting and identifying mycobacterial species
XX Claim 45; Page 76; 91pp; English.
XX This sequence represents a mycobacterial rpoB gene fragment, that is
CC amplified using the PCR primers of the invention. The primers form a
CC method of detecting and identifying mycobacterial species by constructing
CC a phylogenetic tree for the species. The use of the primers for
CC sequence-specific amplification of the rpoB gene (encoding the beta
CC subunit of RNA polymerase) from mycobacterial species provides an
CC efficient way of characterising these species. In addition to
CC phylogenetic analysis, the rpoB gene can be used as an alternative to
CC the 16S rRNA gene because it has four subunits, which are highly
CC conserved throughout prokaryotes. The method is particularly useful for
CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
CC susceptibility can be simultaneously determined in M. tuberculosis.
XX Sequence 306 BP; 51 A; 109 C; 95 G; 51 T; 0 other;
SQ
Query Match 100.0%; Score 20; DB 19; Length 306;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 12
AAAX27219/c
ID AAX27219 standard; DNA; 306 BP.
XX
XX AAX27219;
XX
XX 27-MAY-1999 (first entry)
XX RpoB gene fragment.
XX RpoB gene; mycobacteria; phylogenetic tree construction;
KW mycobacterial species identification; phylogenetic analysis; ss.
XX Nocardia nova.
XX WO9905316-A1.
XX 04-FEB-1999.
XX 28-JUL-1998; 98KR-0000228.

```

```

XX
PR 28-JUL-1997; 97KR-0035501.
XX (BION-) BIONEER CORP.
XX Kim B, Kook Y;
XX WPI; 1998-539367/46.
XX New pair of polymerase chain reaction (PCR) primers - for
PT sequence-specific amplification of the rpoB gene from mycobacterial
PT species, useful for detecting and identifying mycobacterial species
XX Claim 48; Page 77; 91pp; English.
XX This sequence represents a mycobacterial rpoB gene fragment, that is
CC amplified using the PCR primers of the invention. The primers form a
CC method of detecting and identifying mycobacterial species by constructing
CC a phylogenetic tree for the species. The use of the primers for
CC sequence-specific amplification of the rpoB gene (encoding the beta
CC subunit of RNA polymerase) from mycobacterial species provides an
CC efficient way of characterising these species. In addition to
CC phylogenetic analysis, the rpoB gene can be used as an alternative to
CC the 16S rRNA gene because it has four subunits, which are highly
CC conserved throughout prokaryotes. The method is particularly useful for
CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
CC susceptibility can be simultaneously determined in M. tuberculosis.
XX Sequence 306 BP; 53 A; 109 C; 94 G; 50 T; 0 other;
SQ
Query Match 100.0%; Score 20; DB 19; Length 306;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 13
AAAX27220/c
ID AAX27220 standard; DNA; 306 BP.
XX
XX AAX27220;
XX
XX 27-MAY-1999 (first entry)
XX RpoB gene fragment.
XX RpoB gene; mycobacteria; phylogenetic tree construction;
KW mycobacterial species identification; phylogenetic analysis; ss.
XX Rhodococcus equi.
XX WO9905316-A1.
XX 04-FEB-1999.
XX 28-JUL-1998; 98KR-0000228.
XX 28-JUL-1997; 97KR-0035501.
XX (BION-) BIONEER CORP.
XX Kim B, Kook Y;
XX WPI; 1998-539367/46.
XX New pair of polymerase chain reaction (PCR) primers - for
PT sequence-specific amplification of the rpoB gene from mycobacterial
PT species, useful for detecting and identifying mycobacterial species
XX Claim 49; Page 77-78; 91pp; English.

```

XX This sequence represents a mycobacterial rpoB gene fragment, that is
 CC amplified using the PCR primers of the invention. The primers form a
 CC method of detecting and identifying mycobacterial species by constructing
 CC a phylogenetic tree for the species. The use of the primers for
 CC sequence-specific amplification of the rpoB gene (encoding the beta
 CC subunit of RNA polymerase) from mycobacterial species provides an
 CC efficient way of characterising these species. In addition to
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to
 CC the 16S rRNA gene because it has four subunits, which are highly
 CC conserved throughout prokaryotes. The method is particularly useful for
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
 CC susceptibility can be simultaneously determined in M. tuberculosis.
 XX
 SQ Sequence 306 BP; 50 A; 110 C; 98 G; 48 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;
 Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20
 |||||
 Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 14
 AAX27191/c
 ID AAX27191 standard; DNA; 306 BP.
 XX
 AC AAX27191;
 XX
 DT 27-MAY-1999 (first entry)
 XX
 DE RpoB gene fragment.
 XX
 KW RpoB gene; mycobacteria; phylogenetic tree construction;
 KW mycobacterial species identification; phylogenetic analysis; ss.
 XX
 OS Mycobacteria gordoniae.
 XX
 PN WO9905316-A1.
 XX
 PD 04-FEB-1999.
 XX
 PF 28-JUL-1998; 98KR-0000228.
 XX
 PR 28-JUL-1997; 97KR-0035501.
 XX
 PA (BION-) BIONEER CORP.
 XX
 PI Kim B, Kook Y;
 XX
 WP1; 1998-539367/46.

XX New pair of polymerase chain reaction (PCR) primers - for
 PT sequence-specific amplification of the rpoB gene from mycobacterial
 PT species, useful for detecting and identifying mycobacterial species
 XX
 PS Claim 20; Page 68; 91pp; English.
 XX
 CC This sequence represents a mycobacterial rpoB gene fragment, that is
 CC amplified using the PCR primers of the invention. The primers form a
 CC method of detecting and identifying mycobacterial species by constructing
 CC a phylogenetic tree for the species. The use of the primers for
 CC sequence-specific amplification of the rpoB gene (encoding the beta
 CC subunit of RNA polymerase) from mycobacterial species provides an
 CC efficient way of characterising these species. In addition to
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to
 CC the 16S rRNA gene because it has four subunits, which are highly
 CC conserved throughout prokaryotes. The method is particularly useful for
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
 CC susceptibility can be simultaneously determined in M. tuberculosis.
 XX

XX This sequence represents a mycobacterial rpoB gene fragment, that is
 CC amplified using the PCR primers of the invention. The primers form a
 CC method of detecting and identifying mycobacterial species by constructing
 CC a phylogenetic tree for the species. The use of the primers for
 CC sequence-specific amplification of the rpoB gene (encoding the beta
 CC subunit of RNA polymerase) from mycobacterial species provides an
 CC efficient way of characterising these species. In addition to
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to
 CC the 16S rRNA gene because it has four subunits, which are highly
 CC conserved throughout prokaryotes. The method is particularly useful for
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
 CC susceptibility can be simultaneously determined in M. tuberculosis.
 XX

SQ Sequence 306 BP; 52 A; 101 C; 101 G; 52 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;
 Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20
 |||||
 Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 15
 AAX27195/c
 ID AAX27195 standard; DNA; 306 BP.
 XX
 AC AAX27195;
 XX
 DT 27-MAY-1999 (first entry)
 XX
 DE RpoB gene fragment.
 XX
 KW RpoB gene; mycobacteria; phylogenetic tree construction;
 KW mycobacterial species identification; phylogenetic analysis; ss.
 XX
 OS Mycobacteria intracellulare.
 XX
 PN WO9905316-A1.
 XX
 PD 04-FEB-1999.
 XX
 PF 28-JUL-1998; 98KR-0000228.
 XX
 PR 28-JUL-1997; 97KR-0035501.
 XX
 PA (BION-) BIONEER CORP.
 XX
 PI Kim B, Kook Y;
 XX
 WP1; 1998-539367/46.

XX New pair of polymerase chain reaction (PCR) primers - for
 PT sequence-specific amplification of the rpoB gene from mycobacterial
 PT species, useful for detecting and identifying mycobacterial species
 XX
 PS Claim 24; Page 69; 91pp; English.
 XX
 CC This sequence represents a mycobacterial rpoB gene fragment, that is
 CC amplified using the PCR primers of the invention. The primers form a
 CC method of detecting and identifying mycobacterial species by constructing
 CC a phylogenetic tree for the species. The use of the primers for
 CC sequence-specific amplification of the rpoB gene (encoding the beta
 CC subunit of RNA polymerase) from mycobacterial species provides an
 CC efficient way of characterising these species. In addition to
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to
 CC the 16S rRNA gene because it has four subunits, which are highly
 CC conserved throughout prokaryotes. The method is particularly useful for
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
 CC susceptibility can be simultaneously determined in M. tuberculosis.
 XX

XX This sequence represents a mycobacterial rpoB gene fragment, that is
 CC amplified using the PCR primers of the invention. The primers form a
 CC method of detecting and identifying mycobacterial species by constructing
 CC a phylogenetic tree for the species. The use of the primers for
 CC sequence-specific amplification of the rpoB gene (encoding the beta
 CC subunit of RNA polymerase) from mycobacterial species provides an
 CC efficient way of characterising these species. In addition to
 CC phylogenetic analysis, the rpoB gene can be used as an alternative to
 CC the 16S rRNA gene because it has four subunits, which are highly
 CC conserved throughout prokaryotes. The method is particularly useful for
 CC slow growing, fastidious or uncultivable mycobacteria. Also, rifampin
 CC susceptibility can be simultaneously determined in M. tuberculosis.
 XX
 SQ Sequence 306 BP; 51 A; 111 C; 98 G; 46 T; 0 other;

Query Match 100.0%; Score 20; DB 19; Length 306;
 Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGGTCTGC 20
 |||||
 Db 129 GGATGTTGATCAGGGTCTGC 110

Search completed: November 12, 2002, 16:50:34
 Job time : 14.0983 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 ; Search time 61.1561 Seconds
(without alignments)
9517.553 Million cell updates/sec

Title: US-09-697-123B-26

Perfect score: 20

Sequence: 1 ggatgttgatcagggtctgc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vl.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	20	100.0	226	6	A47855	A47855 Sequence 69
C 2	20	100.0	254	6	A47823	A47823 Sequence 37
C 3	20	100.0	306	1	AF057449	AF057449 Mycobacte
C 4	20	100.0	306	1	AF057455	AF057455 Mycobacte
C 5	20	100.0	306	1	AF057456	AF057456 Mycobacte
C 6	20	100.0	306	1	AF057457	AF057457 Mycobacte
C 7	20	100.0	306	1	AF057458	AF057458 Mycobacte
C 8	20	100.0	306	1	AF057459	AF057459 Mycobacte
C 9	20	100.0	306	1	AF057460	AF057460 Mycobacte
C 10	20	100.0	306	1	AF057461	AF057461 Mycobacte
C 11	20	100.0	306	1	AF057462	AF057462 Mycobacte
C 12	20	100.0	306	1	AF057463	AF057463 Mycobacte
C 13	20	100.0	306	1	AF057464	AF057464 Mycobacte
C 14	20	100.0	306	1	AF057465	AF057465 Mycobacte
C 15	20	100.0	306	1	AF057467	AF057467 Mycobacte
C 16	20	100.0	306	1	AF057468	AF057468 Mycobacte
C 17	20	100.0	306	1	AF057472	AF057472 Mycobacte
C 18	20	100.0	306	1	AF057477	AF057477 Mycobacte
C 19	20	100.0	306	1	AF057478	AF057478 Mycobacte
C 20	20	100.0	306	1	AF057479	AF057479 Mycobacte
C 21	20	100.0	306	1	AF057480	AF057480 Mycobacte
C 22	20	100.0	306	1	AF057481	AF057481 Mycobacte
C 23	20	100.0	306	1	AF057482	AF057482 Mycobacte
C 24	20	100.0	306	1	AF057483	AF057483 Mycobacte
C 25	20	100.0	306	1	AF057484	AF057484 Mycobacte
C 26	20	100.0	306	1	AF057485	AF057485 Mycobacte
C 27	20	100.0	306	1	AF057487	AF057487 Mycobacte
C 28	20	100.0	306	1	AF057488	AF057488 Mycobacte
C 29	20	100.0	306	1	AF057489	AF057489 Mycobacte
C 30	20	100.0	306	1	AF057490	AF057490 Mycobacte
C 31	20	100.0	306	1	AF057492	AF057492 Mycobacte
C 32	20	100.0	306	1	AF057494	AF057494 Mycobacte
C 33	20	100.0	306	1	AF057495	AF057495 Rhodococc
C 34	20	100.0	306	1	AF173087	AF173087 Mycobacte
C 35	20	100.0	306	6	AR157002	AR157002 Sequence
C 36	20	100.0	306	6	AR157004	AR157004 Sequence
C 37	20	100.0	306	6	AR157005	AR157005 Sequence
C 38	20	100.0	306	6	AR157006	AR157006 Sequence
C 39	20	100.0	306	6	AR157009	AR157009 Sequence
C 40	20	100.0	306	6	AR157010	AR157010 Sequence
C 41	20	100.0	306	6	AR157011	AR157011 Sequence
C 42	20	100.0	306	6	AR157012	AR157012 Sequence
C 43	20	100.0	306	6	AR157013	AR157013 Sequence
C 44	20	100.0	306	6	AR157014	AR157014 Sequence
C 45	20	100.0	306	6	AR157015	AR157015 Sequence

ALIGNMENTS

RESULT 1
A47855/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

A47855
Sequence 69 from Patent WO9533851.
A47855
A47855.1 GI:2301741
unidentified.
unclassified.
1 (bases 1 to 226)
De.B.H., Portael,F., Machtelinckx,L., Jannes,G. and Rossau,R.
METHOD FOR THE DETECTION OF THE ANTIBIOTIC RESISTANCE SPECTRUM OF
MYCOBACTERIUM SPECIES
Patent: WO 9533851-A 69 14-DEC-1995;

226 bp
DNA
linear
PAT 07-MAR-1997

Pred. No. is the number of results predicted by chance to have a

```

INNOGENETICS NV (BE)
Other publication AU 2789695 960104.
FEATURES
  source
    1..226
    /organism="unidentified"
    /db_xref="taxon:32644"
BASE COUNT      40 a   76 c   73 g   37 t
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 226;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGCTCTGC 20
      |||||||
Db 86 GGATGTTGATCAGGCTCTGC 67

RESULT 2
A47823/c
LOCUS
DEFINITION      Sequence 37 from Patent WO9533851.
ACCESSION      A47823
VERSION        A47823.1 GI:2301709
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 254)
  De,B.H., Portels,F., Machtelinckx,L., Jannes,G. and Rossau,R.
  METHOD FOR THE DETECTION OF THE ANTIBIOTIC RESISTANCE SPECTRUM OF
  MYCOBACTERIUM SPECIES
  Patent: WO 9533851-A 37 14-DEC-1995;
JOURNAL
INNOGENETICS NV (BE)
Other publication AU 2789695 960104.
FEATURES
  source
    1..254
    /organism="unidentified"
    /db_xref="taxon:32644"
BASE COUNT      40 a   84 c   90 g   39 t   1 others
ORIGIN

Query Match      100.0%; Score 20; DB 6; Length 254;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGCTCTGC 20
      |||||||
Db 94 GGATGTTGATCAGGCTCTGC 75

RESULT 3
AF057449/c
LOCUS
DEFINITION      Mycobacterium abscessus RNA polymerase beta (rpoB) gene, partial
  cds.
ACCESSION      AF057449
VERSION        AF057449.1 GI:5902485
KEYWORDS
SOURCE
ORGANISM
Mycobacterium abscessus.
Mycobacterium abscessus
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE
  1 (bases 1 to 306)
  Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
  Kim,E.C., Cha,C.Y. and Kook,Y.H.
  Identification of mycobacterial species by comparative sequence
  analysis of the RNA polymerase gene (rpoB)
  J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
JOURNAL
MEDLINE
PUBMED
99262756
10325313
REFERENCE
  2 (bases 1 to 306)
  Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
  Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
  Direct Submission
  Submitted (06-APR-1998) Microbiology, Seoul National University
  College of Medicine, 28 Yongsong-dong, Chongno-gu, Seoul 110-799,
  Korea
  Location/Qualifiers
    1..306
    /organism="Mycobacterium abscessus"
    /strain="CAP97E-03"
    /db_xref="taxon:36809"
    <1..>306
    /gene="rpoB"
    <1..>306
    /gene="rpoB"
    /codon_start=3
    /transl_table=11
    /product="RNA polymerase beta"
    /protein_id="AAD55513.1"
    /db_xref="GI:5902486"
    /translation="RTVGELIQNRVGLSRMERVVRMTTQDVEAITPQTILNIRP
    VVAIKEFFGTSQLSQPMQDNPLSLGTHKRLSALCPGGLTRDAGLEVRDVHPSH"
BASE COUNT      51 a   108 c   96 g   51 t
ORIGIN

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGCTCTGC 20
      |||||||
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 4
AF057455/c
LOCUS
DEFINITION      Mycobacterium asiaticum RNA polymerase beta (rpoB) gene, partial
  cds.
ACCESSION      AF057455
VERSION        AF057455.1 GI:5902497
KEYWORDS
SOURCE
ORGANISM
Mycobacterium asiaticum.
Mycobacterium asiaticum
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE
  1 (bases 1 to 306)
  Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
  Kim,E.C., Cha,C.Y. and Kook,Y.H.
  Identification of mycobacterial species by comparative sequence
  analysis of the RNA polymerase gene (rpoB)
  J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
JOURNAL
MEDLINE
PUBMED
99262756
10325313
REFERENCE
  2 (bases 1 to 306)
  Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
  Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
  Direct Submission
  Submitted (06-APR-1998) Microbiology, Seoul National University
  College of Medicine, 28 Yongsong-dong, Chongno-gu, Seoul 110-799,
  Korea
  Location/Qualifiers
    1..306
    /organism="Mycobacterium asiaticum"
    /strain="ATCC25276"
    /db_xref="ATCC:25276"
    /db_xref="taxon:1790"
    <1..>306
    /gene="rpoB"
    <1..>306
    /gene="rpoB"
    /codon_start=3

```

```

/transl_table=11
/product="RNA polymerase beta"
/protein_id="AAD5519.1"
/db_xref="GI:5902498"
/translation="RTVGELIQNIQIRVGSRMERVRMTTQDVEAITPOTLINIRP
VVAAIKEFFGTQSOLFMDQNNPLSGLTHKRRLSALGPGLSRAGLEVDRVHPSH"
BASE COUNT      51 a 106 c 98 g 51 t
ORIGIN

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGCTCTGC 20
    |||||||
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 5
AF057456/c
LOCUS      306 bp DNA linear BCT 19-JUN-2000
DEFINITION Mycobacterium aurum RNA polymerase beta (rpoB) gene, partial cds.
ACCESSION  AF057456
VERSION     AF057456.2 GI:8572548
KEYWORDS   .
SOURCE     Mycobacterium aurum.
ORGANISM   Mycobacterium aurum
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE  1 (bases 1 to 306)
AUTHORS   Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
           Kim,E.C., Cha,C.Y. and Kook,Y.H.
TITLE     Identification of mycobacterial species by comparative sequence
           analysis of the RNA polymerase gene (rpoB)
JOURNAL   J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
MEDLINE   99262756
PUBMED    10325313
REFERENCE  2 (bases 1 to 306)
AUTHORS   Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
           Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
TITLE     Direct Submission
JOURNAL   Submitted (06-APR-1998) Microbiology, Seoul National University
           College of Medicine, 28 Yungun-dong, Chongno-gu, Seoul 110-799,
           Korea
REFERENCE  3 (bases 1 to 306)
AUTHORS   Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
           Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
TITLE     Direct Submission
JOURNAL   Submitted (19-JUN-2000) Microbiology, Seoul National University
           College of Medicine, 28 Yungun-dong, Chongno-gu, Seoul 110-799,
           Korea
REMARK     Sequence update by submitter
COMMENT    On Jun 19, 2000 this sequence version replaced gi:5902499.
FEATURES   Location/Qualifiers
            source
              1..306
              /organism="Mycobacterium aurum"
              /strain="ATCC23366"
              /db_xref="ATCC:23366"
              /db_xref="taxon:1791"
              <1..>306
              /gene="rpoB"
              <1..>306
              /gene="rpoB"
              /codon_start=3
              /transl_table=11
              /product="RNA polymerase beta"
              /protein_id="AAD5520.2"
              /db_xref="GI:8572549"
              /translation="RTVGELIQNIQIRVGSRMERVRMTTQDVEAITPOTLINIRP
              VVAAIKEFFGTQSOLFMDQNNPLSGLTHKRRLSALGPGLSRAGLEVDRVHSSH"
BASE COUNT      49 a 107 c 98 g 52 t
ORIGIN

```

```

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGCTCTGC 20
    |||||||
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 6
AF057457/c
LOCUS      306 bp DNA linear BCT 17-SEP-1999
DEFINITION Mycobacterium avium RNA polymerase beta (rpoB) gene, partial cds.
ACCESSION  AF057457
VERSION     AF057457.1 GI:5902501
KEYWORDS   .
SOURCE     Mycobacterium avium.
ORGANISM   Mycobacterium avium
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE  1 (bases 1 to 306)
AUTHORS   Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
           Kim,E.C., Cha,C.Y. and Kook,Y.H.
TITLE     Identification of mycobacterial species by comparative sequence
           analysis of the RNA polymerase gene (rpoB)
JOURNAL   J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
MEDLINE   99262756
PUBMED    10325313
REFERENCE  2 (bases 1 to 306)
AUTHORS   Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
           Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
TITLE     Direct Submission
JOURNAL   Submitted (06-APR-1998) Microbiology, Seoul National University
           College of Medicine, 28 Yungun-dong, Chongno-gu, Seoul 110-799,
           Korea
FEATURES   Location/Qualifiers
            source
              1..306
              /organism="Mycobacterium avium"
              /strain="ATCC25291"
              /db_xref="ATCC:25291"
              /db_xref="taxon:1764"
              <1..>306
              /gene="rpoB"
              <1..>306
              /gene="rpoB"
              /codon_start=3
              /transl_table=11
              /product="RNA polymerase beta"
              /protein_id="AAD5521.1"
              /db_xref="GI:5902502"
              /translation="RTVGELIQNIQIRVGSRMERVRMTTQDVEAITPOTLINIRP
              VVAAIKEFFGTQSOLFMDQNNPLSGLTHKRRLSALGPGLSRAGLEVDRVHPSH"
BASE COUNT      50 a 109 c 101 g 46 t
ORIGIN

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGCTCTGC 20
    |||||||
Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 7
AF057458/c
LOCUS      306 bp DNA linear BCT 17-SEP-1999
DEFINITION Mycobacterium celatum strain ATCC51131 RNA polymerase beta (rpoB)
           gene, partial cds.
ACCESSION  AF057458
VERSION     AF057458.1 GI:5902503
KEYWORDS   .
SOURCE     Mycobacterium celatum strain ATCC51131
ORGANISM   Mycobacterium celatum strain ATCC51131
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE  1 (bases 1 to 306)
AUTHORS   Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
           Kim,E.C., Cha,C.Y. and Kook,Y.H.
TITLE     Identification of mycobacterial species by comparative sequence
           analysis of the RNA polymerase gene (rpoB)
JOURNAL   J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
MEDLINE   99262756
PUBMED    10325313
REFERENCE  2 (bases 1 to 306)
AUTHORS   Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
           Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
TITLE     Direct Submission
JOURNAL   Submitted (06-APR-1998) Microbiology, Seoul National University
           College of Medicine, 28 Yungun-dong, Chongno-gu, Seoul 110-799,
           Korea
FEATURES   Location/Qualifiers
            source
              1..306
              /organism="Mycobacterium celatum strain ATCC51131"
              /strain="ATCC51131"
              /db_xref="ATCC:51131"
              /db_xref="taxon:1764"
              <1..>306
              /gene="rpoB"
              <1..>306
              /gene="rpoB"
              /codon_start=3
              /transl_table=11
              /product="RNA polymerase beta"
              /protein_id="AAD5522.1"
              /db_xref="GI:5902504"
              /translation="RTVGELIQNIQIRVGSRMERVRMTTQDVEAITPOTLINIRP
              VVAAIKEFFGTQSOLFMDQNNPLSGLTHKRRLSALGPGLSRAGLEVDRVHPSH"
BASE COUNT      50 a 109 c 101 g 46 t
ORIGIN

```

KEYWORDS
SOURCE
ORGANISM
Mycobacterium celatum.
Mycobacterium celatum
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
MEDLINE
PUBMED
10325313
REFERENCE
AUTHORS
2 (bases 1 to 306)
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
TITLE
JOURNAL
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Yongsong-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES
source
Location/Qualifiers
1..306
/organism="Mycobacterium celatum"
/strain="ATCC51131"
/db_xref="ATCC:51131"
/db_xref="taxon:28045"
<1..>306
/gene="rpoB"
<1..>306
/gene="rpoB"
/codon_start=3
/transl_table=11
/product="RNA polymerase beta"
/protein_id="AAD5522.1"
/db_xref="GI:5902504"
/translation="RTVGLIQNIQIVGSMRERVVEMTODVEAITPOTLINIRP
VVAIKFEFTQSLSQFMDQNNPLSLGTHRRNLALGPGLSRERAGLEVRDVHPSH"
BASE COUNT 55 a 102 c 101 g 48 t
ORIGIN
Query Match 100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 129 GGATGTTGATCAGGCTCTGC 110
RESULT 8
AF057459/c
LOCUS
DEFINITION
Mycobacterium celatum strain ATCC51130 RNA polymerase beta (rpoB)
gene, partial cds.
ACCESSION
AF057459
VERSION
AF057459.1 GI:5902505
KEYWORDS
SOURCE
ORGANISM
Mycobacterium celatum.
Mycobacterium celatum
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
MEDLINE
PUBMED
10325313
REFERENCE
AUTHORS
2 (bases 1 to 306)
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,

Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Yongsong-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES
source
Location/Qualifiers
1..306
/organism="Mycobacterium celatum"
/strain="ATCC51130"
/db_xref="ATCC:51130"
/db_xref="taxon:28045"
<1..>306
/gene="rpoB"
<1..>306
/gene="rpoB"
/codon_start=2
/transl_table=11
/product="RNA polymerase beta"
/protein_id="AAD5523.1"
/db_xref="GI:5902506"
/translation="SYRRRADPEPDGGRYVEDGAGGRADDPGRGRRGDHAADPDQHPHS
GRGHQGVLRHQPALAVHGPQPIVRADPQASPERARPCWSVPGARPGGPRRAPEPL
" BASE COUNT 55 a 102 c 101 g 48 t
ORIGIN
Query Match 100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATGTTGATCAGGCTCTGC 20
|||||
Db 129 GGATGTTGATCAGGCTCTGC 110
RESULT 9
AF057460/c
LOCUS
DEFINITION
Mycobacterium chelonae RNA polymerase beta (rpoB) gene, partial
cds.
ACCESSION
AF057460
VERSION
AF057460.1 GI:5902507
KEYWORDS
SOURCE
ORGANISM
Mycobacterium chelonae.
Mycobacterium chelonae
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
MEDLINE
PUBMED
10325313
REFERENCE
AUTHORS
2 (bases 1 to 306)
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
TITLE
JOURNAL
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Yongsong-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES
source
Location/Qualifiers
1..306
/organism="Mycobacterium chelonae"
/strain="ATCC35749"
/db_xref="ATCC:35749"
/db_xref="taxon:1774"
<1..>306
/gene="rpoB"
<1..>306
/gene="rpoB"

```

/codon_start=3
/transl_table=11
/product="RNA polymerase beta"
/protein_id="AAD55524.1"
/db_xref="GI:5902508"
/translation="RTVGELIQNIRVGLSRMVRVRRMTTQDVEAITPOTLINIRP
VVAAIKEFFGTSQLSQFMDQNNPLSLGTHKRRLSALPGGLTRDRAGLEVRDVPSPH"
BASE COUNT      51 a   105 c   94 g   56 t
ORIGIN

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGGTCTGC 20
|||||
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 10
AF057461/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED
10325313
AUTHORS
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Youngon-dong, Chongno-gu, Seoul 110-799,
Korea

FEATURES
source
Location/Qualifiers
1..306
/organism="Mycobacterium chitae"
/strain="ATCC19627"
/db_xref="ATCC:19627"
/db_xref="taxon:1792"
<1..>306
/gene="rpoB"
<1..>306
/gene="rpoB"
/codon_start=3
/transl_table=11
/product="RNA polymerase beta"
/protein_id="AAD55525.1"
/db_xref="GI:5902510"
/translation="RTVGELIQNIRVGLSRMVRVRRMTTQDVEAITPOTLINIRP
VVAAIKEFFGTSQLSQFMDQNNPLSLGTHKRRLSALPGGLSRAGLEVRDVPSPH"
BASE COUNT      49 a   108 c   98 g   51 t
ORIGIN

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGGTCTGC 20
|||||
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 11
AF057462/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED
10325313
AUTHORS
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Youngon-dong, Chongno-gu, Seoul 110-799,
Korea

FEATURES
source
Location/Qualifiers
1..306
/organism="Mycobacterium fallax"
/strain="ATCC35219"
/db_xref="ATCC:35219"
/db_xref="taxon:1793"
<1..>306
/gene="rpoB"
<1..>306
/gene="rpoB"
/codon_start=3
/transl_table=11
/product="RNA polymerase beta"
/protein_id="AAD55526.1"
/db_xref="GI:5902512"
/translation="RTVGELIQNIRVGLSRMVRVRRMTTQDVEAITPOTLINIRP
VVAAIKEFFGTSQLSQFMDQNNPLSLGTHKRRLSALPGGLSRAGLEVRDVPSPH"
BASE COUNT      50 a   110 c   101 g   45 t
ORIGIN

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGGTCTGC 20
|||||
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 12
AF057463/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,

```

```

RESULT 11
AF057462/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED
10325313
AUTHORS
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Youngon-dong, Chongno-gu, Seoul 110-799,
Korea

FEATURES
source
Location/Qualifiers
1..306
/organism="Mycobacterium fallax"
/strain="ATCC35219"
/db_xref="ATCC:35219"
/db_xref="taxon:1793"
<1..>306
/gene="rpoB"
<1..>306
/gene="rpoB"
/codon_start=3
/transl_table=11
/product="RNA polymerase beta"
/protein_id="AAD55526.1"
/db_xref="GI:5902512"
/translation="RTVGELIQNIRVGLSRMVRVRRMTTQDVEAITPOTLINIRP
VVAAIKEFFGTSQLSQFMDQNNPLSLGTHKRRLSALPGGLSRAGLEVRDVPSPH"
BASE COUNT      50 a   110 c   101 g   45 t
ORIGIN

Query Match      100.0%; Score 20; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATGTTGATCAGGGTCTGC 20
|||||
Db 129 GGATGTTGATCAGGGTCTGC 110

RESULT 12
AF057463/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 (bases 1 to 306)
Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,

```

```

Kim,E.C., Cha,C.Y. and Kook,Y.H.
Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED
10325313
REFERENCE
2 (bases 1 to 306)
Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
Direct Submission
Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Youngon-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES
    source      Location/Qualifiers
                1..306
                /organism="Mycobacterium flavescens"
                /strain="ATCC14474"
                /db_xref="ATCC:14474"
                /db_xref="taxon:1776"
                <1..>306
                /gene="rpoB"
                <1..>306
                /gene="rpoB"
                /codon_start=3
                /transl_table=11
                /product="RNA polymerase beta"
                /protein_id="AAD5527.1"
                /db_xref="GI:5902514"
                /translation="RTVGLIQNIRVGLSRMERYVRMTTQDVEAITPQTLINEIRP
                VVAAIKEFFGTSQLSQFMDQNNPLSLGTHRRLSALPGGLSRERAGLEVRDVHPSH"
BASE COUNT    49 a 106 c 100 g 51 t
ORIGIN
    Query Match      100.0%; Score 20; DB 1; Length 306;
    Best Local Similarity 100.0%; Pred. No. 26;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGTCGTCG 20
    |||||||
Db 129 GGATGTTGATCAGGTCGTCG 110

RESULT 13
AF057464/c
LOCUS
DEFINITION Mycobacterium fortuitum strain ATCC6841 RNA polymerase beta (rpoB)
gene, partial cds.
ACCESSION AF057464
VERSION AF057464.1 GI:5902515
KEYWORDS
SOURCE Mycobacterium fortuitum.
ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE 1 (bases 1 to 306)
AUTHORS Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
TITLE Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED
10325313
REFERENCE 2 (bases 1 to 306)
AUTHORS Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
TITLE Direct Submission
JOURNAL Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Youngon-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES
    source      Location/Qualifiers
                1..306
                /organism="Mycobacterium fortuitum"

```

```

/strain="ATCC6841"
/db_xref="ATCC:6841"
<1..>306
/genes="rpoB"
<1..>306
/genes="rpoB"
/codon_start=3
/transl_table=11
/product="RNA polymerase beta"
/protein_id="AAD5528.1"
/db_xref="GI:5902516"
/translation="RTVGLIQNIRVGLSRMERYVRMTTQDVEAITPQTLINEIRP
VVAAIKEFFGTSQLSQFMDQNNPLSLGTHRRLSALPGGLSRERAGLEVRDVHSH"
BASE COUNT    49 a 107 c 97 g 53 t
ORIGIN
    Query Match      100.0%; Score 20; DB 1; Length 306;
    Best Local Similarity 100.0%; Pred. No. 26;
    Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGTCGTCG 20
    |||||||
Db 129 GGATGTTGATCAGGTCGTCG 110

RESULT 14
AF057465/c
LOCUS
DEFINITION Mycobacterium fortuitum strain ATCC49403 RNA polymerase beta (rpoB)
gene, partial cds.
ACCESSION AF057465
VERSION AF057465.1 GI:5902517
KEYWORDS
SOURCE Mycobacterium fortuitum.
ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE 1 (bases 1 to 306)
AUTHORS Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
Kim,E.C., Cha,C.Y. and Kook,Y.H.
TITLE Identification of mycobacterial species by comparative sequence
analysis of the RNA polymerase gene (rpoB)
J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
99262756
PUBMED
10325313
REFERENCE 2 (bases 1 to 306)
AUTHORS Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
TITLE Direct Submission
JOURNAL Submitted (06-APR-1998) Microbiology, Seoul National University
College of Medicine, 28 Youngon-dong, Chongno-gu, Seoul 110-799,
Korea
FEATURES
    source      Location/Qualifiers
                1..306
                /organism="Mycobacterium fortuitum"
                /strain="ATCC49403"
                /db_xref="ATCC:49403"
                /db_xref="taxon:1766"
                <1..>306
                /gene="rpoB"
                <1..>306
                /gene="rpoB"
                /codon_start=3
                /transl_table=11
                /product="RNA polymerase beta"
                /protein_id="AAD5529.1"
                /db_xref="GI:5902518"
                /translation="RTVGLIQNIRVGLSRMERYVRMTTQDVEAITPQTLINEIRP
                VVAAIKEFFGTSQLSQFMDQNNPLSLGTHRRLSALPGGLSRERAGLEVRDVHSH"
BASE COUNT    49 a 106 c 98 g 53 t
ORIGIN

```


Query Match 100.0%; Score 20; DB 1; Length 306;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGCTCTGC 20
 |||
 Db 129 GGATGTTGATCAGGCTCTGC 110

RESULT 15

AF057467/c
 LOCUS
 DEFINITION Mycobacterium genavense RNA polymerase beta (rpoB) gene, partial cds.
 ACCESSION AF057467
 VERSION AF057467.1 GI:5902521
 KEYWORDS
 SOURCE
 ORGANISM Mycobacterium genavense.
 Mycobacterium genavense
 Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 Actinomycetales; Corynebacterineae; Mycobacteriaceae;
 Mycobacterium.
 REFERENCE 1 (bases 1 to 306)
 Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H., Chae,G.T.,
 Kim,E.C., Cha,C.Y. and Kook,Y.H.
 Identification of mycobacterial species by comparative sequence
 analysis of the RNA polymerase gene (rpoB)
 J. Clin. Microbiol. 37 (6), 1714-1720 (1999)
 99262756
 MEDLINE 10125313
 PUBMED
 REFERENCE 2 (bases 1 to 306)
 Kook,Y.H., Kim,B.J., Lee,S.H., Lyu,M.A., Kim,S.J., Bai,G.H.,
 Kim,S.J., Chae,G.T., Kim,E.J. and Cha,C.Y.
 Direct Submission
 TITLE Submitted (06-APR-1998) Microbiology, Seoul National University
 JOURNAL College of Medicine, 28 Youngon-dong, Chongno-gu, Seoul 110-799,
 Korea

FEATURES
source

Location/Qualifiers
 1..306
 /organism="Mycobacterium genavense"
 /strain="ATCC51233"
 /db_xref="ATCC:51233"
 /db_xref="taxon:36812"
 <1..>306
 /gene="rpoB"
 <1..>306
 /gene="rpoB"
 /codon_start=3
 /transl_table=11
 /product="RNA polymerase beta"
 /protein_id="AAD5531.1"
 /db_xref="GI:5902522"
 /translation="RTVGLIQNQIRVGMSEVVRVRRMTTQDVEAITPOTLINIRP
 VVAAIKEFGTSQSFMDNNPLSGLTHKRRLSALGPGGLSRERAGLEVRDVFESH"
 BASE COUNT 51 a 99 c 105 g 51 t
 ORIGIN

Query Match 100.0%; Score 20; DB 1; Length 306;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATGTTGATCAGGCTCTGC 20
 |||
 Db 129 GGATGTTGATCAGGCTCTGC 110

Search completed: November 13, 2002, 01:28:16
 Job time : 63.8561 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 ; Search time 101.315 Seconds
(without alignments)
3037.202 Million cell updates/sec

Title: US-09-697-123b-25

Perfect score: 19
Sequence: 1 tcaaggagaaagcgtacga 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estlin:*
- 4: em_estmu:*
- 5: em_estcov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_man:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	19	100.0	1282	9	AI770311 42 Mycob
c 2	18	94.7	705	17	AQ575881 nbxb0088A
c 3	17.4	91.6	609	12	BF495134 AT03975.5
c 4	17.4	91.6	647	13	BM427436 pgfzn.BK0
c 5	17.4	91.6	659	12	BF504233 AT05552.5
c 6	17.4	91.6	663	12	BF498268 AT12651.5

7	17.4	91.6	667	9	AJ451042
c 8	17	89.5	558	17	BH785787 fzmb013f0
9	16.4	86.3	249	17	BH424101 BOGOY49PR
c 10	16.4	86.3	265	10	BE517162 WHE0615.H
c 11	16.4	86.3	290	9	AI948950 wq18b12.x
12	16.4	86.3	375	17	BH648296 BOMFY48TF
13	16.4	86.3	432	14	H06751 Y183a07.r1
14	16.4	86.3	437	13	BM375873 EBem06.SQ
c 15	16.4	86.3	441	17	BH717250 BOMPM36TF
16	16.4	86.3	509	12	BG714670 BOM677024
17	16.4	86.3	567	13	B1168616 RE0434.5
18	16.4	86.3	579	13	BM377270 EBem05.SQ
19	16.4	86.3	596	14	BQ465913 HT01A16T
20	16.4	86.3	623	13	BI780229 EBma08.SQ
21	16.4	86.3	677	14	BQ466189 HT01N10T
22	16.4	86.3	692	9	AL598382 DKF2p133b
c 23	16.4	86.3	715	17	AG086517 Pan_trog1
24	16.4	86.3	730	13	BI568344 603292613
25	16.4	86.3	747	12	BG720538 602691684
c 26	16.4	86.3	806	17	BH428934 BOGG249TR
27	16.4	86.3	860	17	BH590938 BOGCC43TF
28	16.4	86.3	863	12	BG368660 HVSME1002
29	16.4	86.3	863	13	BI547968 603189377
30	16.4	86.3	871	9	AL547893 AL547893
31	16.4	86.3	924	13	BI546826 603189726
32	16.4	86.3	947	12	BG365925 HVSME1000
33	16.4	86.3	964	14	BQ934678 AGENCOURT
34	16.4	86.3	999	17	CNS040UL Tetradon
35	16	84.2	212	14	T23183 C2H2-314T H
36	15.8	83.2	127	10	AW787398 945008C12
37	15.8	83.2	168	13	BM517883 k183h03.Y
38	15.8	83.2	172	13	BM516341 k168h12.Y
39	15.8	83.2	227	10	AW144977 rs44c04.Y
40	15.8	83.2	229	10	AW724941 f9a5nm.f1
41	15.8	83.2	229	10	AW724942 f9a5nm.f2
42	15.8	83.2	244	14	BQ441968 3524_l_26
43	15.8	83.2	267	17	BH231978 1006165D0
44	15.8	83.2	280	10	AW718755 17c05nm.f
45	15.8	83.2	280	10	AW718756 17c05nm.f

ALIGNMENTS

RESULT 1
AI770311/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

AI770311 1282 bp mRNA linear EST 24-JAN-2000
42 Mycobacterium anaerobic stationary phase library Mycobacterium
smegmatis cDNA, mRNA sequence.
AI770311 GI:6742680
EST.
Mycobacterium smegmatis.
Mycobacterium smegmatis
Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium.
1 (bases 1 to 1282)
Murugasu-Oei, B., Tay, A. and Dick, T.
Upregulation of stress response genes and ABC transporters in
anaerobic stationary-phase Mycobacterium smegmatis
Mol. Gen. Genet. 262 (4-5), 677-682 (1999)
20092472
Contact: Murugasu-Oei, B.
Mycobacterium Laboratory
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore 117609, Republic of Singapore
Tel: 65 874 3011
Fax: 65 779 1117
Email: mcbom@imcb.nus.edu.sg
Insert Length: 1282 Std Error: 0.00
Seq primer: T3 Forward; T7 Backward.

FEATURES
Location/Qualifiers
1..1282
source

```

/organism="Mycobacterium smegmatis"
/strain="mc2155"
/db_xref="taxon:1772"
/clone_lib="Mycobacterium anaerobic stationary phase
library"
/lab_host="E. coli XL1-Blue MRP'"
/notes="Vector: Lambda ZAP II; Bacilli were disrupted using
a Mini bead beater (Biospec). RNA was isolated using the
RNeasy protocol (Qiagen). Purified RNA was subjected to 2
rounds of digestion with RNase-free DNase I (Promega).
DNase I was heat-inactivated at 75oC for 5 min. and
removed by using RNeasy columns followed by phenol
extraction and ethanol precipitation. The RNA
preparations were confirmed to be free of genomic DNA
contamination by carrying out PCR and RT-PCR using the
Access kit (Promega) and primers specific for the
histone-like protein gene hlp (Lee et al., 1998). cDNA
was synthesized using random hexamer primers (Promega)
and Stratagene's cDNA synthesis kit. cDNA fragments were
ligated into lambda ZAP II vector and packaged in vitro
using Stratagene reagents."
BASE COUNT      230 a 395 c 405 g 252 t
ORIGIN
Query Match      100.0%; Score 19; DB 9; Length 1282;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 236 TCAAGGAGAGCGCTACGA 218

RESULT 2
AQ575881/c
LOCUS      705 bp DNA linear GSS 02-JUN-1999
DEFINITION nbxb0088A06r CUG1 Rice BAC Library Oryza sativa genomic clone
ACCESSION  nbxb0088A06r, DNA sequence.
VERSION     AQ575881
KEYWORDS    AQ575881.1 GI:4976366
SOURCE      GSS.
ORGANISM    Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 705)
Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: GGAACAGCTATGACCATG
Class: BAC ends
High quality sequence start: 62
High quality sequence stop: 336.
Location/Qualifiers
1. 705
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone_lib="nbxb0088A06r"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/notes="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
HindIII; Rice is one of two most popular grains in the

```

```

world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Arumuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 Kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9 %. Two high density filters,
each containing 18,432 clones (doubly spotted), represent
the whole library for colony screening."
BASE COUNT      190 a 172 c 158 g 181 t 4 others
ORIGIN
Query Match      94.7%; Score 18; DB 17; Length 705;
Best Local Similarity 100.0%; Pred. No. 2,4e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACG 18
|||||
Db 94 TCAAGGAGAGCGCTACG 77

RESULT 3
BF495134/c
LOCUS      609 bp mRNA linear EST 19-APR-2001
DEFINITION AT03975.5prime AT Drosophila melanogaster adult testes pOTB7
Drosophila melanogaster cDNA clone AT03975 5 similar to Med:
FBan0001775 'transcription factor' located on: 3R 100D1-100D1;:
04/07/2001, mRNA sequence.
ACCESSION  BF495134
VERSION     BF495134.2 GI:13685621
KEYWORDS    EST.
SOURCE      fruit fly.
ORGANISM    Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 609)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E., Berman
, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V., Farfan
, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P.,
Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nunoo, J., Pacלב, J.,
Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E.,
Ceiniker, S. and Rubin, G.M.
BDGP/HHMI AT Drosophila EST Project
Unpublished (2000)
On Dec 6, 2000 this sequence version replaced gi:11578435.
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003779; arm:3R [27296749,27528812]
estimated-cyto:100D1-100E3: 04/07/2001
Plate: AT.39 row: G column: 3
High quality sequence stop: 595.
Location/Qualifiers
1. 609
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="AT03975"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates

```

```

AT.121-AT.319; DH5-alpha Tona"
/notes="Organ: ADULT testes; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."
BASE COUNT      120 a   167 c   188 g   134 t
ORIGIN
Query Match      91.6%; Score 17.4; DB 12; Length 609;
Best Local Similarity 94.7%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGGAGACGGCTACGA 19
|||||
Db 159 TCAAGGAGACGGCGACGA 141

RESULT 4
BM427436
LOCUS
DEFINITION
pgf2n.pk006.m12 Normalized Chicken Abdominal Fat Library (pgf2n)
Gallus gallus cDNA clone pgf2n.pk006.m12 5' similar to gi|14091760
ref|NP_114461.1| APP-binding protein 1 [Rattus norvegicus]
gb|AAD09247.1| (U90829) APP-binding protein 1 [Rattus norvegicus],
mRNA sequence.
BM427436
VERSION
BM427436.1 GI:18432610
KEYWORDS
EST.
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 647)
Cogburn,L.A., Morgan,R. and Burnside,J.
ESTs from Normalized Chicken fat cDNA library-USDA/IFAFS Animal
Genome Project
Unpublished (2002)
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES
source
1..647
/organism="Gallus gallus"
/strain="Commercial broiler, Ottawa Research Centre,
leghorn"
/db_xref="taxon:9031"
/clone="pgf2n.pk006.m12"
/clone_lib="Normalized Chicken Abdominal Fat Library
(pgf2n)"
/sex="Male and Female"
/tissue_type="Abdominal Fat"
/dev_stage="Embryonic (d18,d19); post-hatch (d1.w3,w7,w9
,w16,1yr)"
/lab_host="E. coli EMDH10B"
/notes="Vector: pCMVSPORT6; Library made from equivalent
pools of total RNA isolated from each developmental age
(across strains); Single pass sequencing from 5'-end"
BASE COUNT      177 a   133 c   159 g   165 t   13 others
ORIGIN

Query Match      91.6%; Score 17.4; DB 13; Length 647;
Best Local Similarity 94.7%; Pred. No. 4.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGGAGACGGCTACGA 19
|||||
Db 40 TCAAGGAGACGGCGCTACGA 58

AT.121-AT.319; DH5-alpha Tona"
/notes="Organ: ADULT testes; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."
BASE COUNT      120 a   167 c   188 g   134 t
ORIGIN
Query Match      91.6%; Score 17.4; DB 12; Length 609;
Best Local Similarity 94.7%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGGAGACGGCTACGA 19
|||||
Db 159 TCAAGGAGACGGCGACGA 141

RESULT 5
BF504233/c
LOCUS
DEFINITION
BF504233
659 bp mRNA linear EST 19-APR-2001
AT05552.5prime AT Drosophila melanogaster adult testes pOTB7
Drosophila melanogaster cDNA clone AT05552 5 similar to Med:
FBan0001775 'transcription factor' located on: 3R 100D1-100D1;;
04/07/2001, mRNA sequence.
BF504233
VERSION
BF504233.2 GI:13686863
KEYWORDS
EST.
SOURCE
fruit fly.
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 659)
Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Baxter,E., Berman
,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V., Farfan
,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P.,
Liao,G., Miranda,A., Misra,S., Mungall,C.J., Nunoo,J., Pacleb,J.,
Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E.,
Celnikier,S. and Rubin,G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)
On Dec 6, 2000 this sequence version replaced gi:11587534.
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003779: arm:3R [27296749,27528812]
estimated-cyto:100D1-100E3: 04/07/2001
Plate: AT.55 row: E column: 4
High quality sequence stop: 657.
Location/Qualifiers
1..659
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="AT05552"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120; DH5-alpha. Plates
AT.121-AT.319; DH5-alpha Tona"
/notes="Organ: ADULT testes; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."
BASE COUNT      128 a   177 c   213 g   141 t
ORIGIN

Query Match      91.6%; Score 17.4; DB 12; Length 659;
Best Local Similarity 94.7%; Pred. No. 4.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGGAGACGGCTACGA 19
|||||
Db 159 TCAAGGAGACGGCGACGA 141

RESULT 6
BF498268/c
LOCUS
DEFINITION
BF498268
663 bp mRNA linear EST 19-APR-2001
AT12651.5prime AT Drosophila melanogaster adult testes pOTB7
Drosophila melanogaster cDNA clone AT12651 5 similar to Med:
FBan0001775 'transcription factor' located on: 3R 100D1-100D1;;
04/09/2001, mRNA sequence.
BF498268
ACCESSION

```

```

VERSION      BF498268.2  GI:13690169
KEYWORDS
SOURCE       fruit fly.
ORGANISM     Drosophila melanogaster

REFERENCE    1 (bases 1 to 663)
AUTHORS      Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E., Berman
              , B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V., Farfan
              , D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P.,
              Liao, G., Miranda, A., Misra, S., Mungall, C.J., Nunoo, J., Pacleb, J.,
              Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.E.,
              Celnikier, S. and Rubin, G.M.
TITLE        BDGP/HMI AT Drosophila EST Project
JOURNAL      Unpublished (2000)
COMMENT      On Dec 6, 2000 this sequence version replaced gi:11581569.
              Contact: Stapleton, M.
              BDGP

              Lawrence Berkeley National Lab
              One Cyclotron Rd. Berkeley, CA 94720, USA
              Fax: 510 486 6798
              Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
              hit genomic AE003779: arm:3R [27296749,27528812]
              estimated-cyto:100D1-100E3: 04/09/2001
              Plate: AT.126 row: E column: 3
              High quality sequence stop: 624.
              Location/Qualifiers
FEATURES     source
              1..663
              /organism="Drosophila melanogaster"
              /db_xref="taxon:7227"
              /clone="AT12651"
              /clone_lib="AT Drosophila melanogaster adult testes potB7"
              /sex="male"
              /dev_stage="0-3 day old Ore-R males"
              /lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
              AT.121-AT.319: DH5-alpha Tona"
              /note="Organ: ADULT testes; Vector: potB7; Site.1: EcoRI;
              Site.2: XhoI; The mRNA for the testis library was made
              from testes and seminal vesicles hand dissected from 0-3
              day old Ore-R males. RNA kindly provided by the lab of
              Margaret Fuller. Sized fractionated cDNAs were directly
              ligated into potB7. Plasmid cDNA library."
BASE COUNT   128 a 178 c 216 g 141 t
ORIGIN
              1..663
              TCAAGGAGAGCGCTACGA 19
              |||||
              159 TCAAGGAGAGCGCGACGA 141

              Query Match      91.6%; Score 17.4; DB 12; Length 663;
              Best Local Similarity 94.7%; Pred. No. 4.6e+02;
              Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19
Db 159 TCAAGGAGAGCGCGACGA 141

RESULT 7
AJ451042
LOCUS        AJ451042 riken1 Gallus gallus cDNA clone 27g6r1, mRNA sequence.
DEFINITION  AJ451042 riken1 Gallus gallus cDNA clone 27g6r1, mRNA sequence.
ACCESSION   AJ451042
VERSION     AJ451042.1 GI:20218263
KEYWORDS    EST.
SOURCE      Chicken.
ORGANISM    Gallus gallus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
            1 (bases 1 to 667)
            Buerstede, J.M.
            Gallus gallus bursal lymphocyte EST
            Unpublished (2002)
            Contact: Buerstede JM
            Cellular Immunology

            Query Match      89.5%; Score 17; DB 17; Length 558;
            Best Local Similarity 100.0%; Pred. No. 6.8e+02;
            Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAGGAGAGCGCTACGA 19
Db 453 AAGGAGAGCGCTACGA 437

```

```

Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.

FEATURES     source
              1..667
              /organism="Gallus gallus"
              /db_xref="taxon:9031"
              /clone="27g6r1"
              /clone_lib="riken1"
              /cell_type="bursal lymphocyte"
              /dev_stage="2-3 weeks old"
              /note="CB inbred strain"
BASE COUNT   186 a 141 c 170 g 169 t
ORIGIN
              1..667
              TCAAGGAGAGCGCTACGA 19
              |||||
              59 TCAAGGAGAGCGCTACGA 77

              Query Match      91.6%; Score 17.4; DB 9; Length 667;
              Best Local Similarity 94.7%; Pred. No. 4.6e+02;
              Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19
Db 59 TCAAGGAGAGCGCTACGA 77

RESULT 8
BH785787/c
LOCUS        BH785787 558 bp DNA linear GSS 28-MAR-2002
DEFINITION  fzm013f050e10k0 fzm013f050e10 5', DNA sequence.
ACCESSION   BH785787
VERSION     BH785787.1 GI:19790828
KEYWORDS    GSS.
SOURCE      Zea mays.
ORGANISM    Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade: Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 558)
            Budiman, M.A., Freese, R.G., Bedell, J.A., Nunberg, A.N. and Lakey, N.D.
            Genefresher methylation filtered genomic sequences from maize
            Unpublished (2002)
            Contact: Bedell JA
            Orion Genomics, LLC
            4041 Forest Park Ave, St. Louis, MO 63108, USA
            Tel: 314 615 6979
            Fax: 314 615 5975
            Email: jbedell@oriongenomics.com
            Plate: fzm013f050 row: e column: 10
            Seq primer: SK reverse
            Class: shotgun
            High quality sequence stop: 558.
            Location/Qualifiers
FEATURES     source
              1..558
              /organism="Zea mays"
              /cultivar="MO17"
              /db_xref="taxon:4577"
              /clone="fzm013f050e10"
              /clone_lib="fzm013f050 filtered library"
              /note="Organ: leaf; Vector: pBCSK(-); Site.1: HincII; DNA
              prepared from purified nuclei was randomly sheared,
              end-repaired, size fractionated to enrich for the 0.5 to
              5 kb fraction, ligated into HincII-digested pBCSK(-)
              vector and electroporated into E. coli cells."
BASE COUNT   134 a 124 c 141 g 159 t
ORIGIN
              1..558
              TCAAGGAGAGCGCTACGA 19
              |||||
              453 AAGGAGAGCGCTACGA 437

```

```

RESULT 9
BH424101
LOCUS
DEFINITION
BOGOY49TR BOGO Brassica oleracea genomic clone BOGOY49, DNA
sequence.
ACCESSION
BH424101
VERSION
BH424101.1 GI:17609829
KEYWORDS
GSS.
SOURCE
Brassica oleracea.
ORGANISM
Brassica oleracea.
REFERENCE
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 249)
Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other_GSSs: BOGOY49TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: 7R
Class: sheared ends.
BASE COUNT
77 a 47 c 56 g 59 t
ORIGIN
1..249
/organism="Brassica oleracea"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOGOY49"
/clone_lib="BOGO"
/note="Vector: pHD1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHD1 using BstXI linkers"

Query Match 86.3%; Score 16.4; DB 17; Length 249;
Best Local Similarity 94.4%; Pred. No. 9.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGGTACGA 19
|||||
Db 60 CAAGGAGAGCGGTACGA 77

RESULT 10
BES17162/c
LOCUS
DEFINITION
WHE0615_H12_P232A Wheat ABA-treated embryo cDNA library Triticum
aestivum cDNA clone WHE0615_H12_P23, mRNA sequence.
ACCESSION
BES17162
VERSION
BES17162.1 GI:9741192
KEYWORDS
EST.
SOURCE
bread wheat.
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
1 (bases 1 to 255)
Anderson,O.D., Chao,S., Han,P.S., Hsia,C.C., Johnson,R.R., Kang,Y.,
Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L., Tong,J.C., Verhey
,S.D. and Walker-Simmons,M.K.
The structure and function of the expressed portion of the wheat
genomes - ABA-treated embryo library
Unpublished (2000)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center

```

```

800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanders@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Clontech Matchmaker 3' AD primer.
FEATURES
Location/Qualifiers
source
1..265
/organism="Triticum aestivum"
/cultivar="Brevor (soft, white, winter, common wheat)"
/db_xref="taxon:4565"
/clone="WHE0615_H12_P23"
/clone_lib="Wheat ABA-treated embryo cDNA library"
/dev_stage="Seed embryo"
/lab_host="E. coli DH12S"
/note="Vector: pCAD10; Site_1: EcoRI; Site_2: XhoI;
Embryos were cut from mature, dormant seeds and imbedded in
25 microm ABA (abscisic acid) in 5 mM Mes buffer, pH 5.7,
for 12 hr at 22 C. The tissue, total RNA, and poly(A)
RNA were prepared by Steven Verhey in M.K.
Walker-Simmons's lab (USDA-ARS, Washington State Univ.,
Pullman, Washington 99164-6420. A cDNA library was made
by Clontech using a combination of random and oligo dt
primers. Library was plated and archived by Russell
Johnson (Colby College, ME/Walker-Simmons' lab). Plasmid
DNA preparations and DNA sequencing were performed in the
OD Anderson lab (all other authors)."
BASE COUNT
16 a 101 c 87 g 60 t
ORIGIN
1..265
Query Match 86.3%; Score 16.4; DB 10; Length 265;
Best Local Similarity 94.4%; Pred. No. 1e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGGTACGA 19
|||||
Db 169 CAAGGAGAGCGGTACGA 152

RESULT 11
AI948950/c
LOCUS
DEFINITION
wq18b12.x1 NCI-CCAP_Kid12 Homo sapiens cDNA clone IMAGE:2471615 3'
similar to SW:GLI4_HUMAN P10075 GLI4 PROTEIN ;, mRNA sequence.
ACCESSION
AI948950
VERSION
AI948950.1 GI:5741260
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 290)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicqap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 880 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 273.
FEATURES
Location/Qualifiers
source
1..290

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2471615"
/clone_lib="NCI_CGAP_Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid5 was prepared, and ss circles were made in vitro. Following HAP hybridization, this DNA was used as tracer in a subtractive library reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 35 a 101 c 96 g 58 t
ORIGIN
Query Match 86.3%; Score 16.4; DB 9; Length 290;
Best Local Similarity 94.4%; Pred. No. 1e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CAAGGAGAGCGCTACGA 19
|||||
Db 19 CAAGGAGAGCGCTACGA 2

RESULT 12
BH648296
LOCUS
DEFINITION
BOMFY48TF BO_2_3_KB Brassica oleracea genomic clone BOMFY48, DNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Brassica oleracea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 375)
Town.C.D., Van Aken.S., Utterback.T. and Fraser.C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
FEATURES
source
1..375
Location/Qualifiers
/organism="Brassica oleracea"
/strain="T0100DH3"
/db_xref="taxon:3712"
/clone="BOMFY48"
/clone_lib="BO_2_3_KB"
/note="Vector: pHOS1; Site.1: BstXI; 2-3 kb sheared genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT 117 a 72 c 92 g 94 t
ORIGIN
Query Match 86.3%; Score 16.4; DB 17; Length 375;
Best Local Similarity 94.4%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CAAGGAGAGCGCTACGA 19
|||||
Db 105 CAAGGAGAGCGCTACGA 122

RESULT 13
H06751
LOCUS
DEFINITION
Y183a07 rl Soares infant brain INIB Homo sapiens cDNA clone IMAGE:44512 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaaskis,E., Waterston,R., Williamson,A., Woldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1388
High quality sequence stops: 265
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1388 Std Error: 0.00
Seq primer: M13Rpl
High quality sequence stop: 265.
FEATURES
source
1..432
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:417053"
/db_xref="taxon:9606"
/clone="IMAGE:44512"
/clone_lib="Soares infant brain INIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lafmid BA; Site.1: Not I; Site.2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTTCGCCGCGCAGGAATTTTCTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lafmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 98 a 101 c 102 g 125 t
ORIGIN
Query Match 86.3%; Score 16.4; DB 14; Length 432;
Best Local Similarity 94.4%; Pred. No. 1.2e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CAAGGAGAGCGCTACGA 19
|||||
Db 335 CAAGGAGAGCGCTACAA 352

RESULT 14
BM375873
LOCUS
DEFINITION
EBem06_SQ004_K13_R embryo, 21 dpa, no treatment, cv Optic, EBem06 Hordeum vulgare cDNA clone EBem06_SQ004_K13 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare
Eukaryota; Eudicotyledons; Poales; Poaceae; Poaceae; Hordeum
BM375873
EST 23-JUL-2002
Hordeum vulgare cDNA clone EBem06_SQ004_K13 5', mRNA sequence.
BASE COUNT 117 a 72 c 92 g 94 t
ORIGIN
Query Match 86.3%; Score 16.4; DB 17; Length 375;
Best Local Similarity 94.4%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CAAGGAGAGCGCTACGA 19
|||||
Db 105 CAAGGAGAGCGCTACGA 122

BASE COUNT 98 a 101 c 102 g 125 t
ORIGIN
Query Match 86.3%; Score 16.4; DB 14; Length 432;
Best Local Similarity 94.4%; Pred. No. 1.2e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CAAGGAGAGCGCTACGA 19
|||||
Db 335 CAAGGAGAGCGCTACAA 352

RESULT 14
BM375873
LOCUS
DEFINITION
EBem06_SQ004_K13_R embryo, 21 dpa, no treatment, cv Optic, EBem06 Hordeum vulgare cDNA clone EBem06_SQ004_K13 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare
Eukaryota; Eudicotyledons; Poales; Poaceae; Poaceae; Hordeum
BM375873
EST 23-JUL-2002
Hordeum vulgare cDNA clone EBem06_SQ004_K13 5', mRNA sequence.
BASE COUNT 117 a 72 c 92 g 94 t
ORIGIN
Query Match 86.3%; Score 16.4; DB 17; Length 375;
Best Local Similarity 94.4%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CAAGGAGAGCGCTACGA 19
|||||
Db 105 CAAGGAGAGCGCTACGA 122

RESULT 13
H06751
LOCUS
DEFINITION
Y183a07 rl Soares infant brain INIB Homo sapiens cDNA clone IMAGE:44512 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaaskis,E., Waterston,R., Williamson,A., Woldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1388
High quality sequence stops: 265
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1388 Std Error: 0.00
Seq primer: M13Rpl
High quality sequence stop: 265.
FEATURES
source
1..432
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="GDB:417053"
/db_xref="taxon:9606"
/clone="IMAGE:44512"
/clone_lib="Soares infant brain INIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lafmid BA; Site.1: Not I; Site.2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTTCGCCGCGCAGGAATTTTCTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lafmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 98 a 101 c 102 g 125 t
ORIGIN
Query Match 86.3%; Score 16.4; DB 14; Length 432;
Best Local Similarity 94.4%; Pred. No. 1.2e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CAAGGAGAGCGCTACGA 19
|||||
Db 335 CAAGGAGAGCGCTACAA 352

RESULT 14
BM375873
LOCUS
DEFINITION
EBem06_SQ004_K13_R embryo, 21 dpa, no treatment, cv Optic, EBem06 Hordeum vulgare cDNA clone EBem06_SQ004_K13 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare
Eukaryota; Eudicotyledons; Poales; Poaceae; Poaceae; Hordeum
BM375873
EST 23-JUL-2002
Hordeum vulgare cDNA clone EBem06_SQ004_K13 5', mRNA sequence.
BASE COUNT 117 a 72 c 92 g 94 t
ORIGIN
Query Match 86.3%; Score 16.4; DB 17; Length 375;
Best Local Similarity 94.4%; Pred. No. 1.1e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CAAGGAGAGCGCTACGA 19
|||||
Db 105 CAAGGAGAGCGCTACGA 122

KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 437)
AUTHORS Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
TITLE Development of Barley Transcriptome Resources
JOURNAL Unpublished (2001)
COMMENT On Jan 10, 2002 this sequence version replaced gi:18119263.
Contact: Waugh R, Marshall DF
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: est@scri.sari.ac.uk
All sequence has a Phred quality score of 20 or over
Seq primer: M13 reverse.

FEATURES Location/Qualifiers
source
1..437
/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
/clone="EBem06_SQ004_K13"
/clone_lib="embryo, 2l DPA, no treatment, cv Optic,
EBem06"
/tissue_type="embryo"
/dev_stage="2l DPA"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from embryos dissected from developing grains (21
days post anthesis) in glasshouse grown barley plants.
Developed as part of the barley transcriptome resources of
BBSRC/SEERAD funded cereal IGF (Investigating Gene
Function) project."

BASE COUNT 97 a 138 c 154 g 48 t
ORIGIN

Query Match 86.3%; Score 16.4; DB 13; Length 437;
Best Local Similarity 94.4%; Pred. No. 1.2e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTACGA 19
||||||| |
Db 345 CAAGGAGAGCGCTACGA 362

RESULT 15
BH717250/c
LOCUS BH717250 441 bp DNA linear GSS 20-FEB-2002
DEFINITION BOMFM36TF BO_2_3_KB Brassica oleracea genomic clone BOMFM36, DNA
sequence.
ACCESSION BH717250
VERSION BH717250.1 GI:18814189
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 441)
AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other_GSSs: BOMFM36Tr
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523

Fax: 301-838-0208
Email: cdownt@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.

FEATURES Location/Qualifiers
source
1..441
/organism="Brassica oleracea"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOMFM36"
/clone_lib="BO_2_3_KB"
/note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"

BASE COUNT 101 a 111 c 86 g 143 t
ORIGIN

Query Match 86.3%; Score 16.4; DB 17; Length 441;
Best Local Similarity 94.4%; Pred. No. 1.2e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTACGA 19
||||||| |
Db 303 CAAGGAGAGCGCTACGA 286

Search completed: November 13, 2002, 04:01:02
Job time : 114.815 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 ; Search time 1.93295 Seconds
(without alignments)
3487.380 Million cell updates/sec

Title: US-09-697-123B-25

Perfect score: 19
Sequence: 1 tcaagagagcgtacga 19

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	84.2	212	10	US-09-833-381-29
2	15.8	83.2	265	10	US-09-294-093B-4362
3	15.8	83.2	281	10	US-09-294-093B-5081
4	15.8	83.2	1176	10	US-09-815-242-9077
5	15.8	83.2	6082	10	US-09-880-107-2432
6	15.4	81.1	331	10	US-09-860-352-9053
7	15.4	81.1	370	10	US-09-960-352-6002
8	15.4	81.1	375	10	US-09-960-352-9968
9	15.4	81.1	375	10	US-09-960-352-12378
10	15.4	81.1	384	10	US-09-960-352-8128
11	15.4	81.1	386	10	US-09-960-352-11656
12	15.4	81.1	390	10	US-09-960-352-5030
13	15.4	81.1	405	10	US-09-960-352-8824
14	15.4	81.1	414	10	US-09-960-352-14725
15	15.4	81.1	422	10	US-09-960-352-4589
16	15.4	81.1	1072	10	US-09-853-386-105
17	15.4	81.1	1230	10	US-09-960-288-1
18	15	78.9	1238	9	US-09-992-598-356
19	15	78.9	1238	10	US-09-989-722-356

20	15	78.9	1238	10	US-09-989-723-356	Sequence 356, App
21	15	78.9	1238	10	US-09-989-279-356	Sequence 356, App
22	15	78.9	1238	10	US-09-989-727-356	Sequence 356, App
23	15	78.9	1238	10	US-09-989-731-356	Sequence 356, App
24	15	78.9	1238	10	US-09-989-732-356	Sequence 356, App
25	15	78.9	1238	10	US-09-991-073-356	Sequence 356, App
26	15	78.9	1238	10	US-09-990-442-356	Sequence 356, App
27	15	78.9	1238	10	US-09-991-163-356	Sequence 356, App
28	15	78.9	1238	10	US-09-993-604-356	Sequence 356, App
29	15	78.9	1238	10	US-09-989-456-356	Sequence 356, App
30	15	78.9	1238	10	US-09-989-721-356	Sequence 356, App
c 31	14.8	77.9	143	10	US-09-983-965-2238	Sequence 2238, App
32	14.8	77.9	207	10	US-09-294-093B-3785	Sequence 3785, App
33	14.8	77.9	513	10	US-09-815-242-7608	Sequence 7608, App
34	14.8	77.9	520	10	US-09-867-701-6244	Sequence 6244, App
35	14.4	75.8	276	10	US-09-294-093B-4114	Sequence 4114, App
36	14.4	75.8	278	10	US-09-878-574-249	Sequence 249, App
37	14.4	75.8	289	10	US-09-960-352-3361	Sequence 3361, App
38	14.4	75.8	363	10	US-09-974-300-4785	Sequence 4785, App
c 39	14.2	74.7	278	10	US-09-923-876-4313	Sequence 4313, App
40	14.2	74.7	283	10	US-09-294-093B-6201	Sequence 6201, App
41	14.2	74.7	298	10	US-09-294-093B-4897	Sequence 4897, App
42	14.2	74.7	401	10	US-09-867-701-3765	Sequence 3765, App
43	14.2	74.7	415	10	US-09-960-352-1970	Sequence 1970, App
44	14.2	74.7	437	10	US-09-764-870-87	Sequence 87, Appl
45	14.2	74.7	588	10	US-09-867-550-1203	Sequence 1203, App

ALIGNMENTS

RESULT 1
US-09-833-381-29
; Sequence 29, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 212
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(212)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-29

Query Match 84.2%; Score 16; DB 10; Length 212;
Best Local Similarity 84.2%; Pred. No. 20;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGA 19
DB 75 TCANGAGAGCNCCTACGA 93

RESULT 2
US-09-294-093B-4362
; Sequence 4362, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL

```
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 4362
; LENGTH: 265
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700353887H1
; NAME/KEY: unsure
; LOCATION: 69, 75, 80, 87, 206
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-4362

Query Match      83.2%; Score 15.8; DB 10; Length 265;
Best Local Similarity 89.5%; Pred. No. 25;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 TCAAGGAGAGCGCTACGA 19
        |||||
Db      140 TCAAGGAGAGCGCTCCGA 158

RESULT 3
US-09-294-093B-5081
; Sequence 5081, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 5081
; LENGTH: 281
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700355754H1
; NAME/KEY: unsure
; LOCATION: 178, 205, 215
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-5081

Query Match      83.2%; Score 15.8; DB 10; Length 281;
Best Local Similarity 89.5%; Pred. No. 25;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 TCAAGGAGAGCGCTACGA 19
        |||||
Db      243 TCAAGGAGAGCGCTCCGA 261

RESULT 4
US-09-815-242-9077
; Sequence 9077, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
```

```
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9077
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1176)
US-09-815-242-9077

Query Match      83.2%; Score 15.8; DB 10; Length 1176;
Best Local Similarity 89.5%; Pred. No. 29;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 TCAAGGAGAGCGCTACGA 19
        |||||
Db      131 TCAAGGAGAGCGCTACGA 149

RESULT 5
US-09-880-107-2432/c
; Sequence 2432, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2432
; LENGTH: 6082
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M96233
US-09-880-107-2432

Query Match      83.2%; Score 15.8; DB 10; Length 6082;
Best Local Similarity 89.5%; Pred. No. 34;
```

```
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAAGCGCTACGA 19
      ||||| ||||| ||| ||
Db 2645 TCAAGGAGAAGCGCTACGA 2627

RESULT 6
US-09-960-352-9053/c
; Sequence 9053, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 9053
; LENGTH: 331
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 39-LIB3057-003-Q1-K1-B4
US-09-960-352-9053

Query Match 81.1%; Score 15.4; DB 10; Length 331;
Best Local Similarity 94.1%; Pred. No. 41;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAAGCGCTACG 18
      ||||| ||||| |||||
Db 327 CAAGGAGAAGCGCTACG 311

RESULT 7
US-09-960-352-6002/c
; Sequence 6002, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6002
; LENGTH: 365
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 26-LIB3057-001-Q1-K1-G9
US-09-960-352-6002

Query Match 81.1%; Score 15.4; DB 10; Length 365;
Best Local Similarity 94.1%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAAGCGCTACG 18
      ||||| ||||| |||||
Db 309 CAAGGAGAAGCGCTACG 293

RESULT 8
US-09-960-352-9968/c
; Sequence 9968, Application US/09960352
; Patent No. US20020137139A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 9968
; LENGTH: 370
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 43-LIB3057-006-Q1-K1-C12
US-09-960-352-9968

Query Match 81.1%; Score 15.4; DB 10; Length 370;
Best Local Similarity 94.1%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAAGCGCTACG 18
      ||||| ||||| |||||
Db 321 CAAGGAGAAGCGCTACG 305

RESULT 9
US-09-960-352-12378/c
; Sequence 12378, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12378
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 53-LIB3057-014-Q1-K1-F10
US-09-960-352-12378

Query Match 81.1%; Score 15.4; DB 10; Length 375;
Best Local Similarity 94.1%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAAGCGCTACG 18
      ||||| ||||| |||||
Db 319 CAAGGAGAAGCGCTACG 303

RESULT 10
US-09-960-352-8128/c
; Sequence 8128, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
```

; SEQ ID NO 8128
; LENGTH: 384
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (9)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 35-LIB3057-006-Q1-K1-A12
US-09-960-352-8128

Query Match 81.1%; Score 15.4; DB 10; Length 384;
Best Local Similarity 94.1%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACG 18
||||| |||||||
DB 321 CAAGGACAAGCGCTACG 305

RESULT 11

; Sequence 11656, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11656
; LENGTH: 386
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 50-LIB3057-003-Q1-K1-E10
US-09-960-352-11656

Query Match 81.1%; Score 15.4; DB 10; Length 386;
Best Local Similarity 94.1%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACG 18
||||| |||||||
DB 201 CAAGGACAAGCGCTACG 217

RESULT 12

; Sequence 5030, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 5030
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 22-LIB3057-005-Q1-K1-F9
US-09-960-352-5030

Query Match 81.1%; Score 15.4; DB 10; Length 390;
Best Local Similarity 94.1%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACG 18
||||| |||||||
DB 298 CAAGGACAAGCGCTACG 314

RESULT 13

; Sequence 8824, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 8824
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (383),(386),(392)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 38-LIB3057-001-Q1-K1-B10
US-09-960-352-8824

Query Match 81.1%; Score 15.4; DB 10; Length 405;
Best Local Similarity 94.1%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACG 18
||||| |||||||
DB 312 CAAGGACAAGCGCTACG 296

RESULT 14

; Sequence 14725, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 14725
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (346)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 63-LIB3057-002-Q1-K1-H4
US-09-960-352-14725

Query Match 81.1%; Score 15.4; DB 10; Length 414;
Best Local Similarity 94.1%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACG 18
||||| |||||||||
Db 167 CAAGGACAAGCGCTACG 183

RESULT 15

US-09-960-352-4589/c
; Sequence 4589, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4589
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-LIB3057-020-Q1-K1-E11
US-09-960-352-4589

Query Match 81.1%; Score 15.4; DB 10; Length 422;
Best Local Similarity 94.1%; Pred. No. 42;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACG 18
||||| |||||||||
Db 312 CAAGGACAAGCGCTACG 296

Search completed: November 12, 2002, 16:59:30
Job time : 4.93295 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:56:36 : Search time 2.54798 Seconds
(without alignments)
2286.857 Million cell updates/sec

Title: US-09-697-123b-25

Perfect score: 19
Sequence: 1 tcaaggagaagcgctacga 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	100.0	970	1	US-08-250-030-1
2	19	100.0	970	5	PCT-US95-06790-1
3	17.4	91.6	3447	2	US-08-313-185-57
4	17.4	91.6	3447	3	US-09-082-614A-57
5	15.8	83.2	1958	4	US-08-961-527-360
6	15.8	83.2	3752	4	US-08-961-527-208
7	15.8	83.2	4854	4	US-08-961-527-178
8	15.8	83.2	6516	4	US-08-961-527-105
9	15.8	83.2	11309	4	US-08-961-527-108
10	14.8	77.9	2395	4	US-09-724-864-24
11	14.8	77.9	6559	4	US-09-234-186-1
12	14.8	77.9	6559	4	US-09-233-527-1
13	14.8	77.9	6560	5	PCT-US93-05651-1
14	14.4	75.8	515	4	US-08-998-416-125
15	14.4	75.8	711	4	US-09-221-017B-476
16	14.2	74.7	22	2	US-08-494-151-6
17	14.2	74.7	756	4	US-08-413-974-3
18	14.2	74.7	756	4	US-08-434-418-3
19	14.2	74.7	756	4	US-08-433-288-3
20	14.2	74.7	756	4	US-08-174-739A-3
21	14.2	74.7	756	4	US-08-434-256-3
22	14.2	74.7	805	4	US-08-858-207A-1
23	14.2	74.7	810	4	US-08-413-974-5
24	14.2	74.7	810	4	US-08-434-418-5
25	14.2	74.7	810	4	US-08-433-288-5
26	14.2	74.7	810	4	US-08-174-739A-5
27	14.2	74.7	810	4	US-08-434-256-5

28	14.2	74.7	812	4	US-09-091-097-7	Sequence 7, Appli
29	14.2	74.7	1049	4	US-08-961-527-317	Sequence 317, App
30	14.2	74.7	1123	1	US-07-971-096-3	Sequence 3, Appli
31	14.2	74.7	1123	1	US-08-175-096-3	Sequence 3, Appli
32	14.2	74.7	2091	4	US-09-134-001C-1459	Sequence 1459, App
33	14.2	74.7	2495	4	US-08-961-527-255	Sequence 255, App
34	14.2	74.7	3240	4	US-09-171-337A-4	Sequence 4, Appli
35	14.2	74.7	4320	2	US-08-494-151-13	Sequence 13, Appli
36	14.2	74.7	4695	2	US-08-231-193A-57	Sequence 57, Appli
37	14.2	74.7	4695	2	US-08-486-273A-57	Sequence 57, Appli
38	14.2	74.7	4695	3	US-08-940-086A-57	Sequence 57, Appli
39	14.2	74.7	4695	4	US-08-940-035A-57	Sequence 57, Appli
40	14.2	74.7	4695	4	US-08-935-105A-57	Sequence 57, Appli
41	14.2	74.7	4695	4	US-09-648-797-57	Sequence 57, Appli
42	14.2	74.7	6171	4	US-08-961-527-37	Sequence 37, Appli
43	14.2	74.7	7650	4	US-09-221-017B-911	Sequence 911, App
44	14.2	74.7	8948	4	US-09-643-597-119	Sequence 119, App
45	14	73.7	7301	4	US-09-816-088-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-250-030-1
: Sequence 1, Application US/08250030
: Patent No. 5643723
: GENERAL INFORMATION:
: APPLICANT: Persing, David H.
: TITLE OF INVENTION: Detection of a Genetic Locus Encoding
: TITLE OF INVENTION: Resistance to Rifampin in Mycobacterial Cultures and in
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Schwegman, Lundberg & Woessner
: STREET: 3500 IDS Center
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/250,030
: FILING DATE: 26-MAY-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Muetting, Ann M.
: REGISTRATION NUMBER: 33,977
: REFERENCE/DOCKET NUMBER: 150.105U51
: TELEPHONE: 612-339-0331
: TELEFAX: 612-339-3061
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 970 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
US-08-250-030-1

Query Match 100.0%; Score 19; DB 1; Length 970;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAAGCGCTACGA 19
Db 26 TCAAGGAGAAGCGCTACGA 44

RESULT 2
PCT-US95-06790-1
; Sequence 1, Application PC/TUS9506790
; GENERAL INFORMATION:
; APPLICANT: Mayo Foundation for Medical Education and Research
; APPLICANT: and Hoffmann-La Roche Inc.
; TITLE OF INVENTION: Detection of a Genetic Locus Encoding
; TITLE OF INVENTION: Resistance to Rifampin
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06790
; FILING DATE: 26-MAY-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Raasch, Kevin W.
; REGISTRATION NUMBER: 35,651
; REFERENCE/DOCKET NUMBER: 150.105W01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
PCT-US95-06790-1

Query Match 100.0%; Score 19; DB 5; Length 970;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGA 19
Db 26 TCAAGGAGAGCGCTACGA 44

RESULT 3
US-08-313-185-57
; Sequence 57, Application US/08313185
; Patent No. 5851763
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying
; APPLICANT: Honore, Nadine
; APPLICANT: Telenti, Amalio
; APPLICANT: Bodmer, Thomas
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
; TITLE OF INVENTION: In Mycobacterium Tuberculosis
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA

ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,185
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356.0068-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-313-185-57

Query Match 91.6%; Score 17.4; DB 2; Length 3447;
Best Local Similarity 94.7%; Pred. No. 8.2;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGA 19
Db 809 TCAAGGAGAGCGCTACGA 827

RESULT 4
US-09-082-614A-57
; Sequence 57, Application US/09082614A
; Patent No. 6124098
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying
; APPLICANT: Honore, Nadine
; APPLICANT: Telenti, Amalio
; APPLICANT: Bodmer, Thomas
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
; TITLE OF INVENTION: In Mycobacterium Tuberculosis
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/082,614A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/313,185
; FILING DATE: 12-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146

; REFERENCE/DOCKET NUMBER: 02356.0068-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-082-614A-57

Query Match 91.6%; Score 17.4; DB 3; Length 3447;
Best Local Similarity 94.7%; Pred. No. 8.2;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 809 TCAAGGAGAGCGCTACGA 827

RESULT 5

US-08-961-527-360
; Sequence 360, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:

; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB340P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 360:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1958 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-08-961-527-360

Query Match 83.2%; Score 15.8; DB 4; Length 1958;
Best Local Similarity 89.5%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 194 TCAAGGAGAGCGCTACGA 212

RESULT 6

US-08-961-527-208
; Sequence 208, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961,527

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB340P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 208:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3752 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-08-961-527-208

Query Match 83.2%; Score 15.8; DB 4; Length 3752;
Best Local Similarity 89.5%; Pred. No. 49;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19
|||||

Db 2773 TCAAGGAGAGCGCTACGA 2791
|||||

RESULT 7

US-08-961-527-178
; Sequence 178, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961,527


```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 2395
; TYPE: DNA
; ORGANISM: Mouse
US-09-724-864-24

Query Match          77.9%; Score 14.8; DB 4; Length 2395;
Best Local Similarity 88.9%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTACGA 19
|||||  |||||  |||||  |||||  |||||
Db 1796 CAAGGAGAACTCTACGA 1779

RESULT 11
US-09-234-186-1
; Sequence 1, Application US/09234186
; Patent No. 6312947
; GENERAL INFORMATION:
; APPLICANT: Hengartner, H. Robert
; TITLE OF INVENTION: IDENTIFICATION AND CHARACTERIZATION OF A
; TITLE OF INVENTION: GENE WHICH PROTECTS CELLS FROM PROGRAMMED CELL DEATH AND
; FILE REFERENCE: 01997/201005
; CURRENT APPLICATION NUMBER: US/09/234,186
; EARLIER FILING DATE: 1999-01-20
; EARLIER APPLICATION NUMBER: 07/898,933
; EARLIER FILING DATE: 1992-06-12
; EARLIER APPLICATION NUMBER: 07/927,681
; EARLIER FILING DATE: 1992-08-10
; EARLIER APPLICATION NUMBER: 08/288,295
; EARLIER FILING DATE: 1994-08-10
; EARLIER APPLICATION NUMBER: 08/801,248
; EARLIER FILING DATE: 1997-02-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6559
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(6559)
; OTHER INFORMATION: n = A,T,C or G
US-09-234-186-1

Query Match          77.9%; Score 14.8; DB 4; Length 6559;
Best Local Similarity 88.9%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTACGA 19
|||||  |||||  |||||  |||||  |||||
Db 5725 CAAGGAGAGTCTAAGA 5742

RESULT 13
PCT-US93-05651-1
; Sequence 1, Application PC/TUS9305651
; GENERAL INFORMATION:
; TITLE OF INVENTION: A Gene Which Prevents Programmed Cell Death
; NUMBER OF SEQUENCES: 5
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05651
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6560 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-05651-1

Query Match          77.9%; Score 14.8; DB 5; Length 6560;
Best Local Similarity 88.9%; Pred. No. 1.6e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTACGA 19
|||||  |||||  |||||  |||||  |||||
Db 5726 CAAGGAGAGTCTAAGA 5743

RESULT 14
US-08-998-416-125/c
; Sequence 125, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgin
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
```

; CITY: Research Triangle Park
; STATE: No 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meligs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: Pf/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 125:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 515 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORGANISM: PAG1046RP
; US-08-998-416-125

Query Match 75.8%; Score 14.4; DB 4; Length 515;
Best Local Similarity 93.8%; Pred. No. 1.9e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTAC 17
||||| |||||
Db 415 CAAGGAGAGCGCTAC 400

RESULT 15
US-09-221-017B-476/c
; Sequence 476, Application US/09221017B
; Patent No. 644799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA: PCT/AU98/01023
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 476:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORYPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...711
; US-09-221-017B-476

Query Match 75.8%; Score 14.4; DB 4; Length 711;
Best Local Similarity 93.8%; Pred. No. 2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CAAGGAGAGCGCTAC 17
||||| ||||| |||||
Db 217 CAAGGAGAGCGCTAC 202

Search completed: November 12, 2002, 21:03:46
Job time : 8.54798 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 ; Search time 12.4434 Seconds
(without alignments)
3438.621 Million cell updates/sec

Title: US-09-697-123b-25

Perfect score: 19

Sequence: 1 tcaaggagaagcgctacga 19

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*

- 1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	100.0	19	22	AAS05225
2	19	100.0	207	22	AAS05204
3	19	100.0	207	22	AAS05212
4	19	100.0	208	22	AAS05201
5	19	100.0	208	22	AAS05202
6	19	100.0	208	22	AAS05203
7	19	100.0	208	22	AAS05205
8	19	100.0	208	22	AAS05206
9	19	100.0	208	22	AAS05208
					Mycobacterium spec
					Mycobacterium gord
					Mycobacterium szul
					Mycobacterium gord
					Mycobacterium gord
					Mycobacterium gord
					Mycobacterium tube
					Mycobacterium terr
					Mycobacterium kans

10	19	100.0	208	22	AAS05210	Mycobacterium ulce
11	19	100.0	208	22	AAS05211	Mycobacterium mari
12	19	100.0	208	22	AAS05215	Mycobacterium aviu
13	19	100.0	208	22	AAS05216	Mycobacterium bowi
14	19	100.0	208	22	AAS05217	Mycobacterium cela
15	19	100.0	208	22	AAS05219	Mycobacterium fort
16	19	100.0	208	22	AAS05222	Mycobacterium afri
17	19	100.0	208	22	AAS05223	Mycobacterium haem
18	19	100.0	208	22	AAS05224	Mycobacterium xeno
19	19	100.0	211	22	AAS05218	Mycobacterium flav
20	19	100.0	214	22	AAS05207	Mycobacterium chel
21	19	100.0	214	22	AAS05214	Mycobacterium mal
22	19	100.0	214	22	AAS05221	Mycobacterium abs
23	19	100.0	223	22	AAS05209	Mycobacterium scro
24	19	100.0	223	22	AAS05213	Mycobacterium gast
25	19	100.0	970	17	AAT09676	Mycobacterium tube
26	19	100.0	3519	22	AAS05176	Mycobacterium tube
27	19	100.0	3534	22	AAS05207	Mycobacterium tube
28	19	100.0	3853	21	AAA74651	Mycobacterium tube
29	19	100.0	3853	21	AAA74651	M. tuberculosis rp
30	17.4	91.6	3239	23	ABL08943	Drosophila melanog
31	17.4	91.6	3239	23	ABL08943	Drosophila melanog
32	17.4	91.6	3447	14	AAQ51532	M. leprae rpoB gene
33	17.4	91.6	4525	23	ABL08942	Drosophila melanog
34	17.4	91.6	6144	23	ABL03424	Drosophila melanog
35	16.4	86.3	407	22	ABA20983	Human nervous syst
36	16.4	86.3	407	22	AAK89467	Human digestive sy
37	16.4	86.3	13644	23	ABL04469	Drosophila melanog
38	16.4	86.3	17549	23	ABL04468	Drosophila melanog
39	15.8	83.2	265	24	ABL74988	Corn tassal-derive
40	15.8	83.2	261	24	ABL75707	Corn tassal-derive
41	15.8	83.2	1040	19	AAV65213	DNA encoding a S.
42	15.8	83.2	1176	23	AAV55440	Streptococcus pneu
43	15.8	83.2	1958	19	AAV52493	Streptococcus pneu
44	15.8	83.2	2102	23	ABL10513	Drosophila melanog
45	15.8	83.2	3752	19	AAV52341	Streptococcus pneu

ALIGNMENTS

RESULT 1

AAS05225

ID AAS05225 standard; DNA; 19 BP.

XX AAS05225;

AC AAS05225;

XX 07-SEP-2001 (first entry)

DT 07-SEP-2001 (first entry)

XX Mycobacterium species rpoB gene fragment PCR primer RPO5'.

DE Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

XX PCR-restriction fragment length polymorphism analysis; PCR primer; ss.

KW Mycobacterium sp.

OS Mycobacterium sp.

XX WO200131061-A1.

PN 03-MAY-2001.

XX 27-OCT-2000; 2000WO-KR01223.

PF 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

PA Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.

DR New DNA fragments from the rpoB gene of mycobacteria, useful for

XX diagnosis and identification of many mycobacterial species by

PT restriction fragment length polymorphism

PT

XX Disclosure; Page 13; 50pp; English.

XX The present sequence for Mycobacterium species rpoB gene fragment PCR
PS primer RPO5' is used with PCR primer RPO3' (AAS05226) to amplify the
CC rpoB gene region from 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

SO Sequence 19 BP; 7 A; 4 C; 6 G; 2 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.5;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAGAGGAGAGCGCTACGA 19
Db 1 TCAGAGGAGAGCGCTACGA 19

RESULT 2
AAS05204
ID AAS05204 standard; DNA; 207 BP.

AC AAS05204;

DT 07-SEP-2001 (first entry)

DE Mycobacterium gordonae type IV rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium gordonae type IV.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism

PS Claim 1; Page 41; 50pp; English.

CC The present sequence for Mycobacterium gordonae type IV rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from

CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

SO Sequence 207 BP; 45 A; 66 C; 66 G; 30 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 207;
Best Local Similarity 100.0%; Pred. No. 7.5;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAGAGGAGAGCGCTACGA 19
Db 1 TCAGAGGAGAGCGCTACGA 19

RESULT 3
AAS05212
ID AAS05212 standard; DNA; 207 BP.

AC AAS05212;

DT 07-SEP-2001 (first entry)

DE Mycobacterium szulgai rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium szulgai.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism

PS Claim 1; Page 43-44; 50pp; English.

CC The present sequence for Mycobacterium szulgai rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to

CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

Sequence 207 BP; 47 A; 63 C; 66 G; 31 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 207;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACGAGAACGCTACGA 19
|||||
Db 1 TCACGAGAACGCTACGA 19

RESULT 4
AAS05201

ID AAS05201 standard; DNA; 208 BP.

AC AAS05201;

DT 07-SEP-2001 (first entry)

DE Mycobacterium gordonae type I rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.
XX

OS Mycobacterium gordonae type I.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism

PS Claim 1; Page 40; 50pp; English.

CC The present sequence for Mycobacterium gordonae type I rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected

CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

Sequence 208 BP; 46 A; 71 C; 65 G; 26 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 208;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACGAGAACGCTACGA 19
|||||
Db 1 TCACGAGAACGCTACGA 19

RESULT 5
AAS05202

ID AAS05202 standard; DNA; 208 BP.

AC AAS05202;

DT 07-SEP-2001 (first entry)

DE Mycobacterium gordonae type II rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.
XX

OS Mycobacterium gordonae type II.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism

PS Claim 1; Page 40; 50pp; English.

CC The present sequence for Mycobacterium gordonae type II rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.


```
RESULT 8
AAS05206
ID AAS05206 standard; DNA: 208 BP.
XX
AC AAS05206;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium terrae rpoB gene fragment.
XX
KM Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium terrae.
XX
PN W0200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000MO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
DR WPI; 2001-300520/31.
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
PS Claim 1; Page 42; 50pp; English.
XX
CC The present sequence for Mycobacterium terrae rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterium species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterium species.
XX
SQ Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
XX
Query Match 100.0%; Score 19; DB 22; Length 208;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TCAAGGAGAGCGCTACGA 19
DB 1 TCAAGGAGAGCGCTACGA 19
XX
RESULT 9
AAS05208
ID AAS05208 standard; DNA: 208 BP.
XX
AC AAS05208;
XX
```

```
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium kansasii rpoB gene fragment.
XX
KM Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium kansasii.
XX
PN W0200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000MO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
DR WPI; 2001-300520/31.
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
PS Claim 1; Page 42; 50pp; English.
XX
CC The present sequence for Mycobacterium kansasii rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterium species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterium species.
XX
SQ Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;
XX
Query Match 100.0%; Score 19; DB 22; Length 208;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TCAAGGAGAGCGCTACGA 19
DB 1 TCAAGGAGAGCGCTACGA 19
XX
RESULT 10
AAS05210
ID AAS05210 standard; DNA: 208 BP.
XX
AC AAS05210;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium ulcerans rpoB gene fragment.
XX
KM Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.
XX
```



```
XX OS Mycobacterium ulcerans.
XX PN WO200131061-A1.
XX PD 03-MAY-2001.
XX PF 27-OCT-2000; 2000MO-KR01223.
XX PR 27-OCT-1999; 99KR-0046795.
XX PA (ERUM-) ERUME BIOTECH CO LTD.
XX PI Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI; 2001-300520/31.
XX DR
XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for
XX PT diagnosis and identification of many mycobacterial species by
XX PT restriction fragment length polymorphism -
XX PS Claim 1; Page 43; 50pp; English.
XX CC The present sequence for Mycobacterium ulcerans rpoB gene
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX CC various Mycobacterial species. These rpoB gene fragments can be used
XX CC in the diagnosis and identification of Mycobacterium species using a
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)
XX CC method. The method comprises obtaining a restriction fragment length
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX CC amplifying and digesting the DNA fragment from the microorganism to
XX CC be identified and comparing the RFLP patterns from the known rpoB gene
XX CC fragments with the unidentified fragment. The rpoB gene fragments
XX CC are useful to identify a wide range of Mycobacterium species, e.g. for
XX CC diagnosis or to obtain epidemiological and pathogenesis information for
XX CC selection of appropriate therapies, including M. tuberculosis, M. leprae
XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR
XX CC required), and can differentiate between many species in a single
XX CC experiment, including those difficult to distinguish by usual biochemical
XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX CC detecting specific Mycobacterial species.
XX SQ Sequence 208 BP; 45 A; 68 C; 67 G; 28 T; 0 other;
XX
XX Query Match 100.0%; Score 19; DB 22; Length 208;
XX Best Local Similarity 100.0%; Pred. No. 7.5;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TCAAGGAGAACGCGCTACGA 19
DB 1 TCAAGGAGAACGCGCTACGA 19
RESULT 11
AAS05211
ID AAS05211 standard; DNA: 208 BP.
XX
XX AAS05211;
XX AC
XX DT 07-SEP-2001 (first entry)
XX DE Mycobacterium marinum rpoB gene fragment.
XX KM Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX KM PCR-restriction fragment length polymorphism analysis; ds.
XX OS Mycobacterium marinum.
XX PN WO200131061-A1.
XX PD 03-MAY-2001.
```

```
XX PF 27-OCT-2000; 2000MO-KR01223.
XX PR 27-OCT-1999; 99KR-0046795.
XX PA (ERUM-) ERUME BIOTECH CO LTD.
XX PI Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI; 2001-300520/31.
XX DR
XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for
XX PT diagnosis and identification of many mycobacterial species by
XX PT restriction fragment length polymorphism -
XX PS Claim 1; Page 43; 50pp; English.
XX CC The present sequence for Mycobacterium marinum rpoB gene
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX CC various Mycobacterial species. These rpoB gene fragments can be used
XX CC in the diagnosis and identification of Mycobacterium species using a
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)
XX CC method. The method comprises obtaining a restriction fragment length
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX CC amplifying and digesting the DNA fragment from the microorganism to
XX CC be identified and comparing the RFLP patterns from the known rpoB gene
XX CC fragments with the unidentified fragment. The rpoB gene fragments
XX CC are useful to identify a wide range of Mycobacterium species, e.g. for
XX CC diagnosis or to obtain epidemiological and pathogenesis information for
XX CC selection of appropriate therapies, including M. tuberculosis, M. leprae
XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR
XX CC required), and can differentiate between many species in a single
XX CC experiment, including those difficult to distinguish by usual biochemical
XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX CC detecting specific Mycobacterial species.
XX SQ Sequence 208 BP; 45 A; 69 C; 68 G; 26 T; 0 other;
```

```
XX
XX Query Match 100.0%; Score 19; DB 22; Length 208;
XX Best Local Similarity 100.0%; Pred. No. 7.5;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TCAAGGAGAACGCGCTACGA 19
DB 1 TCAAGGAGAACGCGCTACGA 19
RESULT 12
AAS05215
ID AAS05215 standard; DNA: 208 BP.
XX
XX AAS05215;
XX AC
XX DT 07-SEP-2001 (first entry)
XX DE Mycobacterium avium rpoB gene fragment.
XX KM Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX KM PCR-restriction fragment length polymorphism analysis; ds.
XX OS Mycobacterium avium.
XX PN WO200131061-A1.
XX PD 03-MAY-2001.
XX PF 27-OCT-2000; 2000MO-KR01223.
XX PR 27-OCT-1999; 99KR-0046795.
XX PA (ERUM-) ERUME BIOTECH CO LTD.
```

```

XX  Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
PI  WPI: 2001-300520/31.
XX
DR  New DNA fragments from the rpoB gene of mycobacteria, useful for
XX  diagnosis and identification of many mycobacterial species by
PT  restriction fragment length polymorphism -
PT
XX  Claim 1: Page 44: 50pp: English.
PS
XX  The present sequence for Mycobacterium avium rpoB gene
CC  fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC  various Mycobacterial species. These rpoB gene fragments can be used
CC  in the diagnosis and identification of Mycobacterium species using a
CC  novel PCR-restriction fragment length polymorphism analysis (PRA)
CC  method. The method comprises obtaining a restriction fragment length
CC  polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC  amplifying and digesting the DNA fragment from the microorganism to
CC  be identified and comparing the RFLP patterns from the known rpoB gene
CC  fragments with the unidentified fragment. The rpoB gene fragments
CC  are useful to identify a wide range of Mycobacterium species, e.g. for
CC  diagnosis or to obtain epidemiological and pathogenesis information for
CC  selection of appropriate therapies, including M. tuberculosis, M. leprae
CC  and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC  with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC  fragments is rapid, precise, simple and cost effective (only 1 PCR
CC  required), and can differentiate between many species in a single
CC  experiment, including those difficult to distinguish by usual biochemical
CC  tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC  detecting specific Mycobacterial species.
XX
SQ  Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;
Query Match      100.0%; Score 19; DB 22; Length 208;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY  1 TCACGAGAGAGCGCTACGA 19
    ||||||||||||||||
DB  1 TCACGAGAGAGCGCTACGA 19
RESULT 13
AAS05216
ID  AAS05216 standard; DNA: 208 BP.
XX
AC  AAS05216;
XX
DT  07-SEP-2001 (first entry)
XX
DE  Mycobacterium bovis rpoB gene fragment.
XX
KM  Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW  PCR-restriction fragment length polymorphism analysis; ds.
XX
OS  Mycobacterium bovis.
XX
PN  WO200131061-A1.
XX
PD  03-MAY-2001.
XX
PF  27-OCT-2000; 2000WO-KR01223.
XX
PR  27-OCT-1999; 99KR-0046795.
XX
PA  (ERUM-) ERUME BIOTECH CO LTD.
XX  Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
PI  WPI: 2001-300520/31.
XX  New DNA fragments from the rpoB gene of mycobacteria, useful for
PT  diagnosis and identification of many mycobacterial species by
PT  restriction fragment length polymorphism -
PT
XX  Claim 1: Page 45: 50pp: English.
PS
XX  The present sequence for Mycobacterium bovis rpoB gene
CC  fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC  various Mycobacterial species. These rpoB gene fragments can be used
CC  in the diagnosis and identification of Mycobacterium species using a
CC  novel PCR-restriction fragment length polymorphism analysis (PRA)
CC  method. The method comprises obtaining a restriction fragment length
CC  polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC  amplifying and digesting the DNA fragment from the microorganism to
CC  be identified and comparing the RFLP patterns from the known rpoB gene
CC  fragments with the unidentified fragment. The rpoB gene fragments
CC  are useful to identify a wide range of Mycobacterium species, e.g. for
CC  diagnosis or to obtain epidemiological and pathogenesis information for
CC  selection of appropriate therapies, including M. tuberculosis, M. leprae
CC  and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC  with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC  fragments is rapid, precise, simple and cost effective (only 1 PCR
CC  required), and can differentiate between many species in a single
CC  experiment, including those difficult to distinguish by usual biochemical
CC  tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC  detecting specific Mycobacterial species.
XX
SQ  Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
Query Match      100.0%; Score 19; DB 22; Length 208;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY  1 TCACGAGAGAGCGCTACGA 19
    ||||||||||||||||
DB  1 TCACGAGAGAGCGCTACGA 19
RESULT 14
AAS05217
ID  AAS05217 standard; DNA: 208 BP.
XX
AC  AAS05217;
XX
DT  07-SEP-2001 (first entry)
XX
DE  Mycobacterium celatum rpoB gene fragment.
XX
KM  Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW  PCR-restriction fragment length polymorphism analysis; ds.
XX
OS  Mycobacterium celatum.
XX
PN  WO200131061-A1.
XX
PD  03-MAY-2001.
XX
PF  27-OCT-2000; 2000WO-KR01223.
XX
PR  27-OCT-1999; 99KR-0046795.
XX
PA  (ERUM-) ERUME BIOTECH CO LTD.
XX  Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
PI  WPI: 2001-300520/31.
XX  New DNA fragments from the rpoB gene of mycobacteria, useful for
PT  diagnosis and identification of many mycobacterial species by
PT  restriction fragment length polymorphism -
PT
XX  Claim 1: Page 45: 50pp: English.
PS
XX  The present sequence for Mycobacterium celatum rpoB gene

```

```

PT  diagnosis and identification of many mycobacterial species by
PT  restriction fragment length polymorphism -
PT
XX  Claim 1: Page 45: 50pp: English.
PS
XX  The present sequence for Mycobacterium bovis rpoB gene
CC  fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC  various Mycobacterial species. These rpoB gene fragments can be used
CC  in the diagnosis and identification of Mycobacterium species using a
CC  novel PCR-restriction fragment length polymorphism analysis (PRA)
CC  method. The method comprises obtaining a restriction fragment length
CC  polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC  amplifying and digesting the DNA fragment from the microorganism to
CC  be identified and comparing the RFLP patterns from the known rpoB gene
CC  fragments with the unidentified fragment. The rpoB gene fragments
CC  are useful to identify a wide range of Mycobacterium species, e.g. for
CC  diagnosis or to obtain epidemiological and pathogenesis information for
CC  selection of appropriate therapies, including M. tuberculosis, M. leprae
CC  and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC  with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC  fragments is rapid, precise, simple and cost effective (only 1 PCR
CC  required), and can differentiate between many species in a single
CC  experiment, including those difficult to distinguish by usual biochemical
CC  tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC  detecting specific Mycobacterial species.
XX
SQ  Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
Query Match      100.0%; Score 19; DB 22; Length 208;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY  1 TCACGAGAGAGCGCTACGA 19
    ||||||||||||||||
DB  1 TCACGAGAGAGCGCTACGA 19
RESULT 14
AAS05217
ID  AAS05217 standard; DNA: 208 BP.
XX
AC  AAS05217;
XX
DT  07-SEP-2001 (first entry)
XX
DE  Mycobacterium celatum rpoB gene fragment.
XX
KM  Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW  PCR-restriction fragment length polymorphism analysis; ds.
XX
OS  Mycobacterium celatum.
XX
PN  WO200131061-A1.
XX
PD  03-MAY-2001.
XX
PF  27-OCT-2000; 2000WO-KR01223.
XX
PR  27-OCT-1999; 99KR-0046795.
XX
PA  (ERUM-) ERUME BIOTECH CO LTD.
XX  Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
PI  WPI: 2001-300520/31.
XX  New DNA fragments from the rpoB gene of mycobacteria, useful for
PT  diagnosis and identification of many mycobacterial species by
PT  restriction fragment length polymorphism -
PT
XX  Claim 1: Page 45: 50pp: English.
PS
XX  The present sequence for Mycobacterium celatum rpoB gene

```

CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 48 A; 71 C; 64 G; 25 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 208;

Best Local Similarity 100.0%; Pred. No. 7.5;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAAGGAGAAAGCGCTACGA 19

|||||

DB 1 TCAAGGAGAAAGCGCTACGA 19

RESULT 15

AAS05219

ID AAS05219 standard; DNA: 208 BP.

AC AAS05219;

XX 07-SEP-2001 (first entry)

DE Mycobacterium fortuitum rpoB gene fragment.

XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

KM PCR-restriction fragment length polymorphism analysis; ds.

XX Mycobacterium fortuitum.

OS WO200131061-A1.

XX 03-MAY-2001.

PD 27-OCT-2000; 2000WO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -

PS Claim 1; Page 46; 50pp; English.

XX The present sequence for Mycobacterium fortuitum rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,

CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 43 A; 69 C; 69 G; 27 T; 0 other;

Query Match 100.0%; Score 19; DB 22; Length 208;

Best Local Similarity 100.0%; Pred. No. 7.5;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAAGGAGAAAGCGCTACGA 19

|||||

DB 1 TCAAGGAGAAAGCGCTACGA 19

Search completed: November 12, 2002, 16:50:33
Job time : 14.4434 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 ; Search time 58.0983 Seconds
(without alignments)
9517.553 Million cell updates/sec

Title: US-09-697-123b-25

Perfect score: 19
Sequence: 1 tcaagagaagcgtacga 19

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb.ba.*
2: gb.htg.*
3: gb.htg.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vl.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vl.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rpd.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	19	100.0	610	1	MTU318818	AJ318818 Mycobacte
2	19	100.0	610	1	MTU318819	AJ318819 Mycobacte
3	19	100.0	616	1	MTU318813	AJ318813 Mycobacte
4	19	100.0	618	1	MTU318815	AJ318815 Mycobacte
5	19	100.0	618	1	MTU318817	AJ318817 Mycobacte
6	19	100.0	633	1	MTU318814	AJ318814 Mycobacte
7	19	100.0	637	1	MTU318816	AJ318816 Mycobacte
8	19	100.0	639	1	MTU318821	AJ318821 Mycobacte
9	19	100.0	970	6	150706	150706 Sequence 1
10	19	100.0	3534	6	AX111339	AX111339 Sequence
11	19	100.0	3752	1	MSU24494	U24494 Mycobacteri
12	19	100.0	3853	1	MTU12205	U12205 Mycobacteri
13	19	100.0	5084	1	MSGRPOB	L27989 Mycobacteri
14	19	100.0	19352	1	AE006964	AE006964 Mycobacte
15	19	100.0	19770	1	MTC1376	295972 Mycobacteri
16	17.4	91.6	2209	3	AF019754	AF019754 Drosophi
17	17.4	91.6	2238	3	AF057162	AF057162 Drosophi
18	17.4	91.6	2849	3	AF041439	AF041439 Drosophi
19	17.4	91.6	3030	3	AF019753	AF019753 Drosophi
20	17.4	91.6	3244	3	AY069496	AY069496 Drosophi
21	17.4	91.6	3253	3	AF039232	AF039232 Drosophi
22	17.4	91.6	3258	3	AF027729	AF027729 Drosophi
23	17.4	91.6	3447	6	AR067447	AR067447 Sequence
24	17.4	91.6	3941	1	AF242549	AF242549 Amycolato
25	17.4	91.6	4391	3	AF039233	AF039233 Drosophi
26	17.4	91.6	11856	1	AE011788	AE011788 Xanthomon
27	17.4	91.6	37617	1	MLB1790G	214314 M.leprae ge
28	17.4	91.6	52259	2	AC017309	AC017309 Drosophi
29	17.4	91.6	124039	2	AC019900	AC019900 Drosophi
30	17.4	91.6	154298	8	AP003504	AP003504 Oryza sat
31	17.4	91.6	159528	3	AC008310	AC008310 Drosophi
32	17.4	91.6	174036	3	AC010214	AC010214 Drosophi
33	17.4	91.6	232064	3	AE003779	AE003779 Drosophi
34	17.4	91.6	253648	3	AE003638	AE003638 Drosophi
35	17.4	91.6	348950	1	MLEPRTN7	AL583923 Mycobacte
36	16.4	86.3	2045	5	FSADB	D55629 Sea lamprey
37	16.4	86.3	3691	9	BC026311	BC026311 Homo sapi
38	16.4	86.3	5647	2	AC014154	AC014154 Drosophi
39	16.4	86.3	10306	1	AE005712	AE005712 Caulobact
40	16.4	86.3	10393	1	AE005136	AE005136 Halobacte
41	16.4	86.3	11021	1	AE004674	AE004674 Pseudomon
42	16.4	86.3	36642	1	AF029344	AF029344 Burkholde
43	16.4	86.3	39896	1	SKK13	AL512667 Streptomy
44	16.4	86.3	41399	2	AC112021	AC112021 Rattus no
45	16.4	86.3	42210	1	SC1C2	AL031124 Streptomy

ALIGNMENTS

RESULT 1
MTU318818
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1415-97.
ACCESSION
AJ318818
VERSION
AJ318818.1 GI:22208412
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
Herrera, L., Jimenez, M.S. and Saez, J.A.

MTU318818
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1415-97.
610 bp
DNA
circular BCT 09-AUG-2002

```

TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
REFERENCE
2 (bases 1 to 610)
Unpublished
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..610
/organism="Mycobacterium tuberculosis"
/isolate="1415-97"
/db_xref="taxon:1773"
1..610
/gene="rpoB"
<1..>610
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87035.1"
/db_xref="GI:22208413"
/translation="LDIYRKLPGEPPTKESAOQLLENLFFKPKRYDLARVGRYKVNK
KLGHLHVGEPITSTLTEDVATIEYLVRLHGGOTTMTVPGGVETDDIDHFGNR
RLRTVGELIQNIRVGMSEVRVRRMTQDVEAITPOTLINIRPVVAAIKEFFGTS
QPSQFMGQNNPLSGLTHKRLSALGPGGLSRERAGLEVRDHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match 100.0%; Score 19; DB 1; Length 610;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 80 TCAAGGAGAGCGCTACGA 98
|||||

RESULT 2
MTU318819
LOCUS
610 bp DNA circular BCT 09-AUG-2002
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1417-97.
ACCESSION
AJ318819
VERSION
1 GI:22208414
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
REFERENCE
2 (bases 1 to 610)
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..610
/organism="Mycobacterium tuberculosis"
/isolate="1417-97"
/db_xref="taxon:1773"
1..610
/gene="rpoB"
<1..>610
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87035.1"
/db_xref="GI:22208413"
/translation="LDIYRKLPGEPPTKESAOQLLENLFFKPKRYDLARVGRYKVNK
KLGHLHVGEPITSTLTEDVATIEYLVRLHGGOTTMTVPGGVETDDIDHFGNR
RLRTVGELIQNIRVGMSEVRVRRMTQDVEAITPOTLINIRPVVAAIKEFFGTS
QPSQFMGQNNPLSGLTHKRLSALGPGGLSRERAGLEVRDHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match 100.0%; Score 19; DB 1; Length 610;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 80 TCAAGGAGAGCGCTACGA 98
|||||

RESULT 3
MTU318813
LOCUS
616 bp DNA circular BCT 09-AUG-2002
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1763-97.
ACCESSION
AJ318813
VERSION
1 GI:22208402
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
REFERENCE
2 (bases 1 to 616)
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..616
/organism="Mycobacterium tuberculosis"
/isolate="1763-97"
/db_xref="taxon:1773"
1..615
/gene="rpoB"
<1..>615
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87030.1"
/db_xref="GI:22208403"
/translation="LDIYRKLPGEPPTKESAOQLLENLFFKPKRYDLARVGRYKVNK
KLGHLHVGEPITSTLTEDVATIEYLVRLHGGOTTMTVPGGVETDDIDHFGNR
RLRTVGELIQNIRVGMSEVRVRRMTQDVEAITPOTLINIRPVVAAIKEFFGTS
QLSQIFMDQNNPLSGLTHKRLSALGPGGLSRERAGLEVRDHP"
BASE COUNT 125 a 191 c 201 g 99 t
ORIGIN
Query Match 100.0%; Score 19; DB 1; Length 616;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAGCGCTACGA 19

```

```

Db      80  TCAAGGAGAGCGCTACGA 98
|||||
RESULT 4
MTU318815
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 2540-97.
ACCESSION
AJ318815
VERSION
RNA polymerase beta subunit; rpoB gene.
KEYWORDS
Mycobacterium tuberculosis.
ORGANISM
Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 618)
Herrera,L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..618
/organism="Mycobacterium tuberculosis"
/isolate="2540-97"
/db_xref="taxon:1773"
gene
1..618
/genes="rpoB"
CDS
<1..>618
/genes="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"
/translation="TDEALLDYRLKRGPEPTKESAOITLLENLFKFKRYDLARVGR
YKVNKGLGHVEPITSTLTEDVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGELIQNIRVGMSEVRMVTQDVEAITPRTLINIRPVAIAIKE
FFGTSQSQFMQNNPLSGLTHKRLLALGPGLSRERAGLEVDRV"
BASE COUNT 124 a 190 c 207 g 97 t
ORIGIN
/organism="Mycobacterium tuberculosis"
/isolate="2540-97"
/db_xref="taxon:1773"
1..618
/genes="rpoB"
<1..>618
/genes="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"
/translation="TDEALLDYRLKRGPEPTKESAOITLLENLFKFKRYDLARVGR
YKVNKGLGHVEPITSTLTEDVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGELIQNIRVGMSEVRMVTQDVEAITPRTLINIRPVAIAIKE
FFGTSQSQFMQNNPLSGLTHKRLLALGPGLSRERAGLEVDRV"
Query Match 100.0%; Score 19; DB 1; Length 618;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 95 TCAAGGAGAGCGCTACGA 113
|||||
RESULT 5
MTU318817
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 2348-98.
ACCESSION
AJ318817
VERSION
RNA polymerase beta subunit; rpoB gene.
KEYWORDS
Mycobacterium tuberculosis.
ORGANISM
Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 633)
Herrera,L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1..633
/genes="rpoB"
gene
1..633
/genes="rpoB"
CDS
<1..>633
/genes="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"
/translation="TDEALLDYRLKRGPEPTKESAOITLLENLFKFKRYDLARVGR
YKVNKGLGHVEPITSTLTEDVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGELIQNIRVGMSEVRMVTQDVEAITPRTLINIRPVAIAIKE
FFGTSQSQFMQNNPLSGLTHKRLLALGPGLSRERAGLEVDRV"
BASE COUNT 124 a 192 c 207 g 95 t
ORIGIN
/organism="Mycobacterium tuberculosis"
/isolate="2348-98"
/db_xref="taxon:1773"
1..618
/genes="rpoB"
<1..>618
/genes="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"
/translation="TDEALLDYRLKRGPEPTKESAOITLLENLFKFKRYDLARVGR
YKVNKGLGHVEPITSTLTEDVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGELIQNIRVGMSEVRMVTQDVEAITPRTLINIRPVAIAIKE
FFGTSQSQFMQNNPLSGLTHKRLLALGPGLSRERAGLEVDRV"
Query Match 100.0%; Score 19; DB 1; Length 618;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 95 TCAAGGAGAGCGCTACGA 113
|||||
RESULT 6
MTU318814
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1058-97.
ACCESSION
AJ318814
VERSION
RNA polymerase beta subunit; rpoB gene.
KEYWORDS
Mycobacterium tuberculosis.
ORGANISM
Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 633)
Herrera,L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1..633
/genes="rpoB"
gene
1..633
/genes="rpoB"
CDS
<1..>633
/genes="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"
/translation="TDEALLDYRLKRGPEPTKESAOITLLENLFKFKRYDLARVGR
YKVNKGLGHVEPITSTLTEDVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGELIQNIRVGMSEVRMVTQDVEAITPRTLINIRPVAIAIKE
FFGTSQSQFMQNNPLSGLTHKRLLALGPGLSRERAGLEVDRV"
BASE COUNT 124 a 192 c 207 g 95 t
ORIGIN
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1..633
/genes="rpoB"
<1..>633
/genes="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"
/translation="TDEALLDYRLKRGPEPTKESAOITLLENLFKFKRYDLARVGR
YKVNKGLGHVEPITSTLTEDVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGELIQNIRVGMSEVRMVTQDVEAITPRTLINIRPVAIAIKE
FFGTSQSQFMQNNPLSGLTHKRLLALGPGLSRERAGLEVDRV"
Query Match 100.0%; Score 19; DB 1; Length 618;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 95 TCAAGGAGAGCGCTACGA 113
|||||
RESULT 7
MTU318817
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 2348-98.
ACCESSION
AJ318817
VERSION
RNA polymerase beta subunit; rpoB gene.
KEYWORDS
Mycobacterium tuberculosis.
ORGANISM
Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 633)
Herrera,L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1..633
/genes="rpoB"
gene
1..633
/genes="rpoB"
CDS
<1..>633
/genes="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"
/translation="TDEALLDYRLKRGPEPTKESAOITLLENLFKFKRYDLARVGR
YKVNKGLGHVEPITSTLTEDVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGELIQNIRVGMSEVRMVTQDVEAITPRTLINIRPVAIAIKE
FFGTSQSQFMQNNPLSGLTHKRLLALGPGLSRERAGLEVDRV"
BASE COUNT 124 a 190 c 207 g 97 t
ORIGIN
/organism="Mycobacterium tuberculosis"
/isolate="2540-97"
/db_xref="taxon:1773"
1..618
/genes="rpoB"
<1..>618
/genes="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"
/translation="TDEALLDYRLKRGPEPTKESAOITLLENLFKFKRYDLARVGR
YKVNKGLGHVEPITSTLTEDVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGELIQNIRVGMSEVRMVTQDVEAITPRTLINIRPVAIAIKE
FFGTSQSQFMQNNPLSGLTHKRLLALGPGLSRERAGLEVDRV"
Query Match 100.0%; Score 19; DB 1; Length 618;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 95 TCAAGGAGAGCGCTACGA 113
|||||
RESULT 8
MTU318817
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 2348-98.
ACCESSION
AJ318817
VERSION
RNA polymerase beta subunit; rpoB gene.
KEYWORDS
Mycobacterium tuberculosis.
ORGANISM
Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 633)
Herrera,L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1..633
/genes="rpoB"
gene
1..633
/genes="rpoB"
CDS
<1..>633
/genes="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"
/translation="TDEALLDYRLKRGPEPTKESAOITLLENLFKFKRYDLARVGR
YKVNKGLGHVEPITSTLTEDVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGELIQNIRVGMSEVRMVTQDVEAITPRTLINIRPVAIAIKE
FFGTSQSQFMQNNPLSGLTHKRLLALGPGLSRERAGLEVDRV"
BASE COUNT 124 a 190 c 207 g 97 t
ORIGIN
/organism="Mycobacterium tuberculosis"
/isolate="2540-97"
/db_xref="taxon:1773"
1..618
/genes="rpoB"
<1..>618
/genes="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"
/translation="TDEALLDYRLKRGPEPTKESAOITLLENLFKFKRYDLARVGR
YKVNKGLGHVEPITSTLTEDVATIEYLVRLHGGQTTMTVPGVGVPEVETDDID
HFGNRLRTVGELIQNIRVGMSEVRMVTQDVEAITPRTLINIRPVAIAIKE
FFGTSQSQFMQNNPLSGLTHKRLLALGPGLSRERAGLEVDRV"
Query Match 100.0%; Score 19; DB 1; Length 618;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 95 TCAAGGAGAGCGCTACGA 113
|||||

```

```

CDS
<1..>633
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87031.1"
/db_xref="GI:22208405"
/translation="DNTVGYDEALLDIYKRLRPGESPPTKESAOTLLENLFFKRYDLARV
ARGRYKVKKGLGHGEPITSTLTEDVVAITYLVRLHGGOTTMTVPGGVEVPEVDD
TDDIHFGNRLRTVGLIQNIQIRVGMSEMRVVRMTTQDVEAITPQTLLINIRPVV
AAIKFEGTSQLSQFMDQNNPLSLGTLKRLSALGPGLSRERAGLEVRDV"
BASE COUNT      129 a   195 c   210 g   99 t
ORIGIN
Query Match      100.0%; Score 19; DB 1; Length 633;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCAAGGAGAGCGGTACGA 19
|||||
Db 110 TCAAGGAGAGCGGTACGA 128

RESULT 7
MTU318816
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1255-98.
ACCESSION
AJ318816
VERSION
AJ318816.1 GI:22208408
KEYWORDS
Mycobacterium tuberculosis.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1 Herrera,L., Jimenez,M.S. and Saez,J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 637)
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..637
/organism="Mycobacterium tuberculosis"
/isolate="1255-98"
/db_xref="taxon:1773"
gene
1..637
/gene="rpoB"
CDS
<1..>637
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87033.1"
/db_xref="GI:22208409"
/translation="VGTDEALLDIYKRLRPGEPPTKESAOTLLENLFFKRYDLARV
GRYKVKKGLGHGEPITSTLTEDVVAITYLVRLHGGOTTMTVPGGVEVPEVDD
IDHFGNRLRTVGLIQNIQIRVGMSEMRVVRMTTQDVEAITPQTLLINIRPVAAI
KEFFGTSQLSQFMDQNNPLSLGTLKRRMFALGPGLSRERAGLEVRDVHPSH"
BASE COUNT      128 a   198 c   210 g   101 t
ORIGIN
Query Match      100.0%; Score 19; DB 1; Length 637;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CDS
<1..>633
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87031.1"
/db_xref="GI:22208405"
/translation="DNTVGYDEALLDIYKRLRPGESPPTKESAOTLLENLFFKRYDL
ARGRYKVKKGLGHGEPITSTLTEDVVAITYLVRLHGGOTTMTVPGGVEVPEV
TDDIHFGNRLRTVGLIQNIQIRVGMSEMRVVRMTTQDVEAITPQTLLINIRPVV
AAIKFEGTSQLSQFMDQNNPLSLGTLKRLSALGPGLSRERAGLEVRDV"
BASE COUNT      129 a   195 c   210 g   99 t
ORIGIN
Query Match      100.0%; Score 19; DB 1; Length 633;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCAAGGAGAGCGGTACGA 19
|||||
Db 101 TCAAGGAGAGCGGTACGA 119

RESULT 8
MTU318821
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1071-98.
ACCESSION
AJ318821
VERSION
AJ318821.1 GI:22208418
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1 Herrera,L., Jimenez,M.S. and Saez,J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 639)
AUTHORS
Herrera,L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..639
/organism="Mycobacterium tuberculosis"
/isolate="1071-98"
/db_xref="taxon:1773"
gene
1..639
/gene="rpoB"
CDS
<1..>639
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87038.1"
/db_xref="GI:22208419"
/translation="RRTDEALLDIYKRLRPGEPPTKESAOTLLENLFFKRYDLARV
GRYKVKKGLGHGEPITSTLTEDVVAITYLVRLHGGOTTMTVPGGVEVPEVDD
IDHFGNRLRTVGLIQNIQIRVGMSEMRVVRMTTQDVEAITPQTLLINIRPVAAI
KEFFGTSQLSQFMDQNNPLSLGTLKRRLSALGPGLSRERAGLEVRDVHPSHS"
BASE COUNT      126 a   202 c   212 g   99 t
ORIGIN
Query Match      100.0%; Score 19; DB 1; Length 639;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TCAAGGAGAGCGGTACGA 19
|||||
Db 101 TCAAGGAGAGCGGTACGA 119

RESULT 9
I50706
LOCUS
DEFINITION
Sequence 1 from patent US 5643723.
ACCESSION
I50706
VERSION
I50706.1 GI:2472409
KEYWORDS
.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 970)
AUTHORS
Persing,D.H., Hunt,J.J., Young,K.K.Y., Felmler,T.A., Roberts,G.D.
and Whelan,A.Christian.

```

TITLE Detection of a genetic locus encoding resistance to rifampin in mycobacterial cultures and in clinical specimens
JOURNAL Patent: US 5643723-A 1 01-JUL-1997;
FEATURES Location/Qualifiers
source 1..970
BASE COUNT 182 a 302 c 330 g 156 t
ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 970;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 26 TCAAGGAGAGCGCTACGA 44

RESULT 10
AX111339
LOCUS
DEFINITION Sequence 2072 from Patent WO0123604.
ACCESSION AX111339
VERSION
KEYWORDS
SOURCE
ORGANISM
Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 3534)
Bergeron, M.G., Boissinot, M., Huletsky, A., m Nard, C., Ouellette, M.,
Picard, F.J. and Roy, P.H.
Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
Patent: WO 0123604-A 2072 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
Location/Qualifiers
source 1..3534
/organism="Mycobacterium tuberculosis"
/strain="Rv"
/db_xref="taxon:1773"
BASE COUNT 679 a 1081 c 1188 g 586 t
ORIGIN

Query Match 100.0%; Score 19; DB 6; Length 3534;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 902 TCAAGGAGAGCGCTACGA 920

RESULT 11
MSU24494
LOCUS
DEFINITION Mycobacterium smegmatis DNA linear BCT 02-MAR-2000
ACCESSION U24494
VERSION U24494.1 GI:790347
KEYWORDS
SOURCE
ORGANISM
Mycobacterium smegmatis.
Mycobacterium smegmatis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
1 (bases 1 to 3752)
Hetherington, S.V., Watson, A.S. and Patrick, C.C.
Sequence and analysis of the rpoB gene of Mycobacterium smegmatis
Antimicrob. Agents Chemother. 39 (9), 2164-2166 (1995)
MEDLINE 96050766
PUBMED 8540740
REFERENCE 2 (bases 1 to 3752)

AUTHORS Hetherington, S.V.
TITLE Direct Submission
JOURNAL Submitted (11-APR-1995) Seth V. Hetherington, Infectious Diseases,
St. Jude Children's Research Hospital, 332 N. Lauderdale, Memphis,
TN 38101, USA
FEATURES Location/Qualifiers
source 1..3752
/organism="Mycobacterium smegmatis"
/db_xref="taxon:1772"
194..3703
/gene="rpoB"
194..3703
/gene="rpoB"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="DNA polymerase"
/protein_id="AAA91426.2"
/db_xref="GI:7144498"
/translation="MLEGCCILAVSSQSKSNAITNNSVRGAPNGVSFAKLREPLEVRGL
LDVOTDSFEMLVGSRVRYRAIDRGEENPVGGLAEVLAELSPIDESGMSLSFSOPR
FDEVKASVDECKDKDNTYAAPLVTAETINNTEIKSQTFVMDGDFPMMEKGTFIIN
GTERVVSQVLRSPGVTFDETDKSTKTLHSVKVIPGRGAWLEFDVDRDVTGVRID
RRKRPVTVLLKALGMTNEQIVERFGSEIMMGTLEKDTTSGTDEALLDIYKLRPGE
YLVRLHREGQTSMTVPVGVFAVESHDIHFGRNRLRTVGLIONQIRVLSRMERVVR
ERMTTODVEALTPTQLINIRPVVAAIKEFTGTSQVSQFMDQNNPLSGLTHKRRLSALG
PGGLSRERAGLEVRDHPHSYHGMCFIETPEGNIGLIGLSVYARVNPFGFTETPYR
KYVGVVDQDYLITADEEDRHVVAQNSPTDENGRTEDRVNVRKKGGEVEFVSADQ
VDMGVSSHKTVTIEVSADYITVMADDTQSRVLRKFRFARNHGTSCANORPIVDA
AARDVAGQVIADGCTQNGEMALKNLLVAIMPGEHNYEDAILLSNRLVEEDVLTISI
HIEHEITDARDTKLGAEEITRVIPKLSDEVLDLSPRIVRIGAEVRDSEILVGVKVP
KGTELTPEERLLRAGFEKAREVDTSLKVPGESGKVGIGRVFSREDDDELUPGVN
ELRVVYAAQRKISDQKLARRHNGVIGKILPVEDMPLPDGTVDILNTHGCRV
VLNIGQLLTHLWGAGKAGNIDVLGAVPDWASKLPPELYSAPADSTVATVPFDGAGE
GELAGLLSTLRDGEVMYNADKATLFDGRSSGPPFPYPTVGVYMLKHLHVDV
KIHARSTGYSMTIQPLGKAQFGGQFGFMECWMAQVGAAYTLQELLTIKSDDTV
GRVYVEAIVKGNIPGPIPESEFKVLLKELQSLCLNVELSSDGRAIEMRDGDDDL
ERAAANLGINLSRNESASVEDLA"
BASE COUNT 706 a 1225 c 1210 g 611 t
ORIGIN

Query Match 100.0%; Score 19; DB 1; Length 3752;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 1068 TCAAGGAGAGCGCTACGA 1086

RESULT 12
MTU12205
LOCUS
DEFINITION Mycobacterium tuberculosis H37Rv RNA-polymerase beta subunit (rpoB)
gene, partial cds.
ACCESSION U12205
VERSION U12205.1 GI:515684
KEYWORDS
SOURCE
ORGANISM
Mycobacterium tuberculosis.
Mycobacterium tuberculosis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 3853)
Imboden, P., Troller, R., Marchesi, F., Telenti, A., Bodmer, T.,
Cole, S., Schopfer, K. and Burkart, T.
The rpoB gene of Mycobacterium tuberculosis
Unpublished
REFERENCE 2 (bases 1 to 3853)
Imboden, P.
Direct Submission

JOURNAL Submitted (11-JUL-1994) Paul Imboden, Institute for Medical Microbiology, University of Berne, Friedbuehlstrasse 51, Berne, 3010, Switzerland

FEATURES
source 1. 3853
Location/Qualifiers
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
576 .>3853
/gene="rpoB"
576 .>3853
/gene="rpoB"
/codon_start=1
/transl_table=1
/product="RNA polymerase beta subunit"
/protein_id="AA20242.2"
/db_xref="GI:7144499"
/translation="MLEGCILADSRQSKTAASPSRPSQSSNNNSVPGAPNRVSFAKL
REPLEVGLDVTQDSFEMWIGSPRMSAAERGVNPGVLEEVLELSPEDFSGS
MSLSFSDPRFDVVKAPVDECKDKMTYAAPLFTVAETNNNTGETKQVPMGDFPMH
TEKGTFLINGTERVVSQVSPGVYFDETIDKSDTKLHSVKVIPSRGAWLEFDVK
RDTGVGRIDRKRQPVTVLLKALGTSQIIVERGFSSEIMRSTLEKONTVGTDEALLD
LYRKLPGEPPTKESAOQLLENLFKERRDLARVGRYKNNKLGHLHVGEPITSSLT
EEDVATIEYLVRLEHQTMTVPGVEVPVETDDIDHFNRRRLTVGELIQNGIRVG
MSRMERVRERTTQDVEAITPQTLINRPVYAAIKEFFGTSQLSQFMDQNNPLSGIT
HKRRSALGPGGLSRERAGLEVDRVHPHYGRMCPETPEGPNIGLIGSLSVYARVNP
FGFIETPYRKYVDGVSDGIVYLTADIEDRHVQAQNSPIDADGRFVEPVLVRKAG
EVEYVPSSEVDYMDVSPROMYSVATAMIPLEHDDANRALGMANQQAQVPLVRSEAP
LVGTGMELRAAIDAATSSQSGVIEEVSADYIIVMDHNGTRRTYRMRKFARNHGT
ANQCPLVDAGDRVEAGQVADGPTDGGEMALGNLLVAIMPWEGHNYEDAIIILSNRL
VEDVLTSHIEHEIDARDTKLGAEEITRIPNISDEVLAIDLDERGIVRIGAEVRDG
DILVGVTPKGETELTPEERLLRAIFGKAREVRDTSIKVPHGESGKVIIGIRVFSRED
EDELPAVGNELRVYVVAQKRKISDGLAGRHGKNGVIGKILPVEDMPFLADGTPVDI
ILNTHGVPFRNNGIQLLETHLGCWAHSGWKVDAAGVPMNAARLPDELEAQAIVS
TPVFGCAEAELOGLSCTLPNRGDVLDVADGKAMLPDGRSGEPFPYPTVGTNTIM
KLHLVDKIHARSTGYPYSMITQPLGKAGQGGQREGMECWAMQAYGAAYTLQELL
TIKS"

BASE COUNT 723 a 1173 c 1293 g 664 t

ORIGIN
Query Match 100.0%; Score 19; DB 1; Length 3853;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 1477 TCAAGGAGAGCGCTACGA 1495

RESULT 13
MSGRPOB 5084 bp DNA linear BCT 13-SEP-1994
LOCUS
DEFINITION
Mycobacterium tuberculosis RNA polymerase beta-subunit (rpoB)
gene, complete cds and RNA polymerase beta'-subunit rpoC gene,
partial cds.
L27989
ACCESSION
VERSION L27989.1 GI:468333
KEYWORDS
SOURCE
ORGANISM
Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium
tuberculosis complex.
1 (bases 1 to 5084)
REFERENCE
AUTHORS Miller,L.P., Crawford,J.T. and Shinnick,T.M.
TITLE The rpoB gene of Mycobacterium tuberculosis
JOURNAL Antimicrob. Agents Chemother. 38 (4), 805-811 (1994)
MEDLINE 94304130
PUBMED
FEATURES
Location/Qualifiers
1. .5084
/organism="Mycobacterium tuberculosis"
/strain="Rv"

gene 1065. .4598
CDS 1065. .4598
/gene="rpoB"
/codon_start=1
/evidence=experimental
/transl_table=1
/product="RNA polymerase beta-subunit"
/protein_id="AA21416.1"
/db_xref="GI:468334"
/translation="MLEGCILADSRQSKTAASPSRPSQSSNNNSVPGAPNRVSFAKL
REPLEVGLDVTQDSFEMWIGSPRMSAAERGVNPGVLEEVLELSPEDFSGS
MSLSFSDPRFDVVKAPVDECKDKMTYAAPLFTVAETNNNTGETKQVPMGDFPMH
TEKGTFLINGTERVVSQVSPGVYFDETIDKSDTKLHSVKVIPSRGAWLEFDVK
RDTGVGRIDRKRQPVTVLLKALGTSQIIVERGFSSEIMRSTLEKONTVGTDEALLD
LYRKLPGEPPTKESAOQLLENLFKERRDLARVGRYKNNKLGHLHVGEPITSSLT
EEDVATIEYLVRLEHQTMTVPGVEVPVETDDIDHFNRRRLTVGELIQNGIRVG
MSRMERVRERTTQDVEAITPQTLINRPVYAAIKEFFGTSQLSQFMDQNNPLSGIT
HKRRSALGPGGLSRERAGLEVDRVHPHYGRMCPETPEGPNIGLIGSLSVYARVNP
FGFIETPYRKYVDGVSDGIVYLTADIEDRHVQAQNSPIDADGRFVEPVLVRKAG
EVEYVPSSEVDYMDVSPROMYSVATAMIPLEHDDANRALGMANQQAQVPLVRSEAP
LVGTGMELRAAIDAATSSQSGVIEEVSADYIIVMDHNGTRRTYRMRKFARNHGT
ANQCPLVDAGDRVEAGQVADGPTDGGEMALGNLLVAIMPWEGHNYEDAIIILSNRL
VEDVLTSHIEHEIDARDTKLGAEEITRIPNISDEVLAIDLDERGIVRIGAEVRDG
DILVGVTPKGETELTPEERLLRAIFGKAREVRDTSIKVPHGESGKVIIGIRVFSRED
EDELPAVGNELRVYVVAQKRKISDGLAGRHGKNGVIGKILPVEDMPFLADGTPVDI
ILNTHGVPFRNNGIQLLETHLGCWAHSGWKVDAAGVPMNAARLPDELEAQAIVS
TPVFGCAEAELOGLSCTLPNRGDVLDVADGKAMLPDGRSGEPFPYPTVGTNTIM
KLHLVDKIHARSTGYPYSMITQPLGKAGQGGQREGMECWAMQAYGAAYTLQELL
TIKS"

BASE COUNT 969 a 1534 c 1691 g 890 t

ORIGIN
Query Match 100.0%; Score 19; DB 1; Length 5084;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 1966 TCAAGGAGAGCGCTACGA 1984

RESULT 14
AE006964
LOCUS
DEFINITION
Mycobacterium tuberculosis CDC1551,
complete genome.
ACCESSION AE006964 AE000516
VERSION AE006964.1 GI:13880217
KEYWORDS
SOURCE
ORGANISM
Mycobacterium tuberculosis CDC1551.
Mycobacterium tuberculosis CDC1551
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium tuberculosis complex.
1 (bases 1 to 19352)
REFERENCE
AUTHORS Fleischmann,R.D., Alland,D., Eisen,J.A., Carpenter,L., White,O.,
Peterson,J., DeBoy,R., Dodson,R., Gwinn,M., Haft,D., Hickey,E.,
Kolony,J.F., Nelson,W.C., Umayam,L.A., Ermolaeva,M.,

TITLE
JOURNAL
AUTHORS
REFERENCE

Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains
Unpublished
2 (bases 1 to 19352)
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
Direct Submission
Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
1..19352
/organism="Mycobacterium tuberculosis CDC1551"
/strain="CDC1551"
/db_xref="taxon:83331"
/note="clinical strain"
163..3699
/gene="MT0695"
163..3699
/gene="MT0695"
/note="similar to GB:L27989 GB:U05910 GB:U12205 SP:P47766
PID:149992; identified by sequence similarity; putative"

FEATURES
source

1..19352
/product="DNA-directed RNA polymerase, beta subunit"
/protein_id="AAK44921.1"
/db_xref="GI:13880218"
/translation="MLEGCIILADSRQSKTAASPSRPOSSNNNSVPGAPNRYSFAL
RELPVGLDVTDSFWMGLSPWRESAARDGNVPGLEEVLYELSPEDFSGS
MSLSDFDFDVKAPVDECKDKMTYAPLVTAETFINNTGKISQTFVMDGFPMM
TKGFIINGTERTVVSVOLVSPGVYFETDKSTKFLHSVKVPSRGAWLEFDVDK
RUTVGRIDRRKQVPIVLLKALGWTSEIIVERFSGSEIMRSTLEKONTVGTDEALID
IRKLPGEPTKSAQTLLENLFKFKERGLDARVGRVYKNNKLGHLVGEPIITSSLT
EDWATLEYLVRLHEGQTTTPVGVGEVPEVETIDHFGNRLRTVGLVGEPIITSSLT
MSRMERVVRRMTQDVEAITPQTLINIRPVVAAIKFEGFTSOLSPMDONNPLSGIT
HKRLSALGPGLSRAGLEVRDVPVHPSHYGRMCP IETPEGNIGLIGLSVYARVNP
FCGFIETPVKVVGVSDIYVLTADIEDRHVQAQNSPIDADGRFVPEVVRKAG
EVEYVPSSEVOYMDVSPQWVSATAMIFLEHODANRGMNQVAPVLRSEAP
LVGTCHERLAAIDAGVVVAEESGVEEVESADYITVMDNGTRRTYRMKFRNSHGT
CANQCPIDVAGDRVAGVGIADGCTDDGEMALGNLLVAIMEGHNYEDAILNSR
LYEEDVLSYIIEHEIDARDKLGAEITRIDIPNIDEVLDLDERGIVRIGAEVRD
GDILGVKTPGKETLTPERLLRAIFGEKAREVDTSLKVPHGESCKVIGIEVSRE
DEDELPAQVNLVRVYVAAKRIISDGDKLAGHCKGVICKILIPVDMPLAQTGTPVD
ITLNVHGPWRNIGQLLETHLGCASGHKVDAAKGVPMWAARLPDELLEAQPNAIV
STPVFDGAQAEALGGLLSLTLPNRDGDVLDADGKAMLFDRSGSEPPFPVTVGYMI
MKLHLVDOKTHARSTGYSMITQOPLGGKAQFGGRFGEKMWQAQYGAAYTLQEL
LTKSDDTVGRVKVYEAIVKGENIPEFGIPESFKVLLKELQSLCLNVLSLSDGAIE
LRGEDEDLERAAANLGINLSRESASVEDIA"
3744..7694
/gene="MT0696"
3744..7694
/gene="MT0696"
/note="similar to SP:P37871; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA-directed RNA polymerase, beta-prime subunit"
/protein_id="AAK44922.1"
/db_xref="GI:13880219"
/translation="MLDVNFFDELRIGLATAEDITROMSYGEVKKPETINVTLPKPKED
GLFCEKIFGPTDRWECYKRVKRFKGIICRCGEVETRVKVRERRRMGHIELAAPT
HIWFKGVPRLGVLDDLPKDLKIIYFAAYVITSVDEMRHNESTLEAEMAVERK
AVEQDQCELEAQAQLEADLAELAEAGAKADARRKVRUGGERMQIKRDAQRELD
LEDIWSFTKLAQQLIVDENLRFVLDVRYGFTGMAGESIQKIENPFDIAEAS
LRDVIRNGQKKLRALKRLKVAQSQSGNSPMGVLDAPVIPPELRPMQLDQGR
FATSDNLVYRRVNRNRLKRLDLAGPEIIVNNEKRLMESVDALFQNGRRPVT
GPNRPLKSLDLKQKGRFQNLKGRVDSGRSVIVVGPOLKHLQCGLPKLMAL
LFKPFVWMLVLDLNAQNTKSAKMYERORPOVMDVLEVIAEHPVLLNRAETLRLG
IQATEPHVEGKALQLHLPLCEAFNDFDGDQMAVHLPLSAEQAQAEARILMLSSNNIL

gene
CDS

gene
CDS

SPASGRPLAMPRLDMVTGLYLLTTEVPDTEYQPASGDHPETGYVSSPAAEIAAADR
GVLSVRKILVRLTDLRPPVETAEELFGHSGWQPGDMAETTLGRVWFNELLPLGY
FVNQMHKKVQAAIINDLAPYPMIVAQTVDKLDAGFYWATRSQVTVSNADVLVPP
RKEILDHYEERADKVEKQFQKALNHDNEALVEIKWKEATDEVGQALREHYPDND
IITIVDSAGTGNFTQTRTLAGMKGLVTPNPKFEIPRPVKSSEFGLTVLEYFINTHGA
RKLADTLARTADSGYTLRRLVDVSDVIREHDDQCTERGIIVVELAERAPDGTILRDP
YIETSAITRLGTDAVEAGNVIVERGQDLGPEQDITALLAAGITQVKRVSVLTGATST
GVATCYGRSMATKLVDTGEAVIVAAQSIGEPGTLTMRTHGGVGEDITGLPR
VOELFEARPGRKAPIADTVGRVLEDEREYKJITIVPDGGEVYVYDKISKRLRV
FKHEDGSERVLSGDHVEVGQQLMEGSDAPHEVLRVQSPREYQJHLVREVQEVTRAQG
VSIHDKHIEVIRQMLRRYIIDSSTEFPLGSLIDRAEFENRVRVVAEGEPPAQR
PLVMGITKASLATDSWLSAASFQETTRVLTDAAINCRSKLNGKLNKENVIIIGKLIPAGT
GNNRYNIAVQPTSEARAAYTIPSYEDQYISPDFGAATGAAPVLDYDYSYR"
/gene="MT0697"
complement(7691..8065)
complement(7691..8065)
/gene="MT0697"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44923.1"
/db_xref="GI:13880220"
/translation="MFSAAALTNPGHAWASAMERSGLLECAGLDQEPFETADKL
NPDGSSRRVPRROADGGIATHVERGGQSGGQAGVYVQPMHGFALMAQDRLIHH
GQOTQNRQAQAFRRVFCVCSPT"
complement(8058..9972)
/gene="MT0698"
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact; identified by
Glimmer2; putative; conserved hypothetical protein,
authentic frameshift"
10167..10925
/gene="MT0699"
10167..10925
/gene="MT0699"
/note="identified by match to PFAM protein family HMM
PF01261"
/codon_start=1
/transl_table=11
/product="AP endonuclease, family 2"
/protein_id="AAK44924.1"
/db_xref="GI:13880221"
/translation="MLGSHVSPDPLAAAEAGDVVQIFLGNPSWKAKPKPRDDAA
ALKATDPIYVHAPYLINLASANNRIPSRKILQETCAAADIGAAAVIVHGHVAD
DNIDKGFQWRKALDRLETPVPLENTAGDHDHARRDITARLMDVIGDTGIGPC
LDTCHTWAAGALTDADVRIKAITGRIDLVCNDRDEAGSGRRHRANLGGQIDPDL
LVAAVKAAGAPVICETADQGRKDDIAFLRRTGS"
10957..11799
/gene="MT0700"
10957..11799
/gene="MT0700"
/note="similar to GB:U00012 PID:466863; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hydrolase/esterase, putative"
/protein_id="AAK44925.1"
/db_xref="GI:13880222"
/translation="MLRVAIILAAVLAFAGCGSGTFLAAGFCNGNSVHTLDVAGAR
SYRLYKVPGLSSAPLVMLHSGKQAKERSYGHDELADSEKFLVAVPDGYHRAWN
ANGGCGCRPARPEGVDIDGFVRAVDIANNVSIIDPARVITVGHNSGAITSYTLACN
SIFAIVGSGTQLDPCQSPRPVSIHIHGTADPLVRYHGGPGAGFARIDGPPVLDLN
AFREVNRCGALDITTEGPTTSGATCADNRNRVLLTVDDAGHRMPSFATQTLWRFEA
AHFR"
11859..13487
/gene="MT0701"
11859..13487
/gene="MT0701"
/note="similar to SP:P33224 GB:L20915 PID:457172
PID:457174 PID:537028; identified by sequence similarity;
putative"
/codon_start=1

gene
CDS

gene
CDS

```

/transl_table=11
/product="acyl-CoA dehydrogenase, putative"
/protein_id="AAK44926.1"
/db_xref="GI:13880223"
/transl_table="MSDTHVTNQVPPLENNYPASSPVLIEALIQEGGQGLDEYNEV
GATISACQARWELADNRNPILTHDAYGYRDEVEYDPAYHELMRTAITHGMAAP
WADRPQHVRAKTSVMTVEPGHICPISMTYAVVVPALRYNSELAAYVEPLLTSEY
DPELPATTKAGLSCTAGMTKEGSDGDRAGTTOATPNADGYSLTGHKWFSAQPCD
IFLVAQAPDGLSCFLPVLDPDTRNMFLOPKDKLGNHANASSEVEYDCAWLV
GEGRGVPTIEMVNLTRLDALGATSMTGLTRAVHHAQHRKAFGAYLIDQPLMRN
VLADLAVEAATIAVMRMAGATDNVAGNETEALLRRIGLAAAKYWCKRSTAAAE
ALECLGNGVSDSGMPRLYREAPLNGIWEGSCNVSLDTRAMATRPACVEVLFDEL
ARSAGDPRLDGHVERLRPQLGLDTIGYRARKIAEDICLALQGLSLVVRHGHPAFAE
FLATRLGGWGGGAYTMPAGLDLAPILERALVKG"
13498..14436
/gene="MT0702"
CDS
13498..14436
/gene="MT0702"
/notes="similar to GP:3885480; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="enoyl-CoA hydratase/isomerase family protein"
/protein_id="AAK44927.1"
/db_xref="GI:13880224"
/transl_table="MTHAIRPVDNLKMTYEVGTGRIARITFNRPEKGNIAIDTPL
ELSAVERADLDGVHVLVSGRGEFCAGFDLSAAGSSSTGGGAYQGTVLDDGT
QAVNHLPNQPMWPMIDYQNMRFVRGFAISMADKPTVVKIHGYCVAGTDIALHADQ
VIAAADAKIGYPTPRVGVGPAAGLWAHRLGDQRKLLFTGDCITGAQAQAWGLAEE
PEPADIDTERLVARIALPVNLIMWLKLSALLQGVATSRMVSIVFDGAARHT
PEGHAFVADAEVGHFRDARRRDEPFDDYGRASRV"
14439..15161
/gene="MT0703"
CDS
14439..15161
/gene="MT0703"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
100.0%; Score 19; DB 1; Length 19352;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGA 19
|||||
Db 1064 TCAAGGAGAGCGCTACGA 1082

RESULT 15
LOCUS MTC1376 19770 bp DNA linear BCT 03-AUG-2001
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 37/162.
ACCESSION 295972 AL123456
VERSION 295972.1 GI:3261790
KEYWORDS
SOURCE Mycobacterium tuberculosis H37Rv.
ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 19770)
Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,
Harris,D., Gordon,S.V., Eiglmeier,K., Gas,S., Barry III,C.E.,
Tekala,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,
Connor,R., Davies,R., Devlin,K., Feltwell,T., Gentles,S.,
Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J.,
Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A.,
Rajandream,M.A., Rogers,J., Rutter,S., Seeger,K., Skelton,S.,
Squires,S., Squares,R., Sulston,J.E., Taylor,K., Whitehead,S., and
Barrell,B.G.
Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence

```

Nature 393 (6685), 537-544 (1998)
98295987
6934230
2 (bases 1 to 19770)
Parkhill,J.
Direct Submission
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 27, 1998 this sequence version replaced gi:2143285.
Notes:
Details of M. tuberculosis sequencing at the Sanger Centre are available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/M.tuberculosis/) CDS have been renumbered from the original cosmid submissions but the Old gene designations are in brackets after the new gene numbers.
Gene prediction was based on a Hidden Markov Model of TB genes implemented in TBPase (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C.
CAUTION: In some cases we may not have predicted the correct (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.
Location/Qualifiers
1..19770
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
/clone_xref="I376"
57..61
/note="agga, possible rbs upstream of Rv0654"
68..1573
/gene="Rv0654"
68..1573
/gene="Rv0654"
/note="Rv0654"
/note="Rv0654", (MTC1376.22), len: 501. unknown, FASTA score: Q53353 LIGNOSTILBENE-ALPHA, BETA-DIOXYGENASE (485 aa) opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity in 523 aa overlap). Also similar to M. tuberculosis protein MTCY21C12.07c (29.5% identity in 522 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0654"
/protein_id="CAB09380.1"
/db_xref="GI:2143307"
/db_xref="SPTREMBL:O06785"
/translation="MTTAAAESQNPYLEGLAPVSTEVATDLPVTGRIPHELDGRY LRNGPNVAEVDPATYTHWFTGDAMVHGVALRDGKARWYRNRVTPAVCAALGEPISA RPHRTGIEGPNVTLTHAGRTLALVAGVWVVELTDELTVGCDPDTLHGVT AHQRDPHTGELHAVSYSFARGHRYQVSVICTGDGHARTVDIEVAGSPMHSFLTDN YLYVLDVPTDPMQVPSVPRWLQRPALVQSVLGRVRIPIPIAALNRHQGHSD RYVYANPSYPARYGVMPREGNEDVRFWFDIEPCYVYTHPLNAYSECRNGAEVLVQW RYSRNFDRDRGPGGDSRLDRTINLATGAVTAECRDDRAEFFPRINETLVGGPHR FAYTVIEGGLVAGAAALSTPLYKQCVTCGTSVASLDPLLLIGENHVFVNPSPARAE DDGILMGYGHWRGDEGOLLDDAQTLLESTIATVHLQRPVPMGFHGNWPTT"
1585..2664
/gene="Rv0655"
1585..2664
/gene="Rv0655"
/note="Rv0655", (MTC1376.21), len: 359. abc transporter, FASTA score: YRBF_ECOLI P45393 hypothetical abc transporter atp-binding (269 aa) opt: 644 z-score: 721.8 E(): 3.4e-33 (38.5% identity in 244 aa overlap); contains PS00017 ATP/GTP-binding site motif A, PS00211 ABC transporters family signature, highly similar to M. leprae MKL_MYCLE P30769 possible ribonucleotide transport atp- (347 aa) opt: 2021 z-score:2244.4 E(): 0, (92.2% identity in 335 aa overlap). Also similar to many otherM.

```

tuberculosis ABC transporters eg. MTCY253.24 (33.6%
identity in 241 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0655"
/protein_id="CAB09379.1"
/db_xref="GI:2143306"
/db_xref="SWISS-PROT:O06784"
/translation="MRYSDSYHTGRQPRASTEGFPMGVSVIEVNGLTAKSFGSSRIWE
DYLTLPAGEVSVLLGPSGCKSVFLKSLGLLRPERGSIIDGTDIIKESAKELYEI
RFLFVGLFQDGLGSMNLYDNTAPFLREHTKKSEIRDIVMEKIALVGLGDKERKF
PGEISGGRKAGLARALVLDPIILDCDEPDGLDPVRYATYLSQLIMDINAQIDATIL
IYTHNINARTVPNMGMLFRKHLVMEFGPEVRLTSDPVVROFLNGRRIGPIGMSEE
KDEATMAEQALLDAGHHAGGVEIEGVPPQISATPCMPERKAVARQARVREMLHTL
PKKAQAAILDDLETKHYAVHICQ"
misc_feature
1762..1785
/gene="Rv0655"
/note="PS00017 ATP/GTP-binding site motif A"
2074..2118
/gene="Rv0655"
/note="PS00211 ABC transporters family signature"
complement(3052..3435)
/gene="Rv0656c"
complement(3052..3435)
/gene="Rv0656c"
/note="Rv0656c"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0656c"
/protein_id="CAB09378.1"
/db_xref="GI:2143305"
/db_xref="SPTREMBL:O06783"
/translation="MAAATTGTHRGELRAAORAVGSCPEPAAEFCSRARNADFEQD
MSRMFGVDPDPVPKSVKSWRWDISAQHRRLARAGALSVLLDLCDAARAGLVVLH
DDADYELAEHRLEDIRRVVSADD"
complement(3530..3685)
/gene="Rv0657c"
complement(3530..3685)
/gene="Rv0657c"
/note="Rv0657c"
/translation="MSVTQIDLDLADVMRTAAVHTKKEAVNLAMRDYVERFRRIE
ALARSRE"
complement(3691..3696)
/note="possible RBS upstream of Rv0657c"
complement(3761..4477)
/gene="Rv0658c"
complement(3761..4477)
/gene="Rv0658c"
/note="Rv0658c"
/translation="MTCI376.18), len: 238, unknown,
probablemembrane protein, similar to YPRB_ECOLI_P33774
hypothetical24.3 kd protein (urf 1) (217 aa), fasta
scores: opt: 174 z-score: 215.8 E(): 5.3e-05, (25.6%
identity in 223 aa overlap). Also similar to MTCY359.10
(28.7% identity in 178 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0658c"
/protein_id="CAB09376.1"
/db_xref="GI:2143303"
/db_xref="SPTREMBL:O06781"
/translation="MEAGRADTVAPSHRWGLGALVWELVFLVASTSLAVVLTGHGPV

```

```

SAGVLALALAAPVVAAGLAILITRLRNGNLTDLRLRNSRGLRLGLMFGFGMLVT
IPASLVYTAIVGPEANSVAVRIFGCVASWPMALVFLVVFVAPLCEEIYRGLLWG
AVDI:RWGRWAALVVTVTVFALAHLEFARAPLLVVAIPIALARYSGLLASIVTHQV
TNLLPGIIVLLGLTGAISLP"
complement(4480..4483)
/note="possible RBS upstream of Rv0658c"
complement(4753..5061)
/gene="Rv0659c"
complement(4753..5061)
/note="Rv0659c"
/translation="MTCI376.17), len: 102; unknown, similar
to YW28_MYCTU_Q10867 hypothetical 12.3 kd protein cy39.28
(114 aa), fasta scores: opt: 144 z-score: 213.2 E():
7.3e-05, 30.8% identity in 107 aa overlap. Also similar to
MTCY09F9.22 (32.7% identity in 101 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0659c"
/protein_id="CAB09387.1"
/db_xref="GI:2143302"
/db_xref="SPTREMBL:O06780"
/translation="MRRGELWFAATPGGDRPVLVLTDRDPVADRIGAVVVALTRTRRG
LVSELELTAVENRVPSDCVYVNFNIIHTLPRTAFRRITRLSPARLHEACQTLRASTGC
"
complement(5048..5293)
/gene="Rv0660c"
complement(5048..5293)
/gene="Rv0660c"
/translation="MTCI376.16), len: 81, some similarity to
IAF016485.130 Halobacterium sp; NRC-1 plasm (100 aa),
32.4% identity in 74 aa overlap"
/codon_start=1

```

```

Query Match      100.0%; Score 19; DB 1; Length 19770;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGA 19
      |||||
DB 10758 TCAGGAGAGCGCTACGA 10776

Search completed: November 13, 2002, 01:28:14
Job time : 73.7983 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 ; Search time 1109.13 Seconds
(Without alignments)
3037.202 Million cell updates/sec

Title: US-09-697-123B-22

Perfect score: 208
Sequence: 1 tcaaggaagcgcctacgac.....ccggtgaacacgcagacat 208

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estb:*
2: em_esth:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estol:*
7: em_estro:*
8: em_estc:*
9: em_estc1:*
10: em_estc2:*
11: em_estc3:*
12: em_estc4:*
13: em_estc5:*
14: em_estc6:*
15: em_estc7:*
16: em_estc8:*
17: em_estc9:*
18: em_estc10:*
19: em_estc11:*
20: em_estc12:*
21: em_estc13:*
22: em_estc14:*
23: em_estc15:*
24: em_estc16:*
25: em_estc17:*
26: em_estc18:*
27: em_estc19:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	150.4	72.3	1282	9	AT170311 42 Mycoba
C 2	43.4	20.9	877	17	A0687771 nxb0075L
C 3	42.2	20.3	568	13	B1995346 B1995346
C 4	39	18.8	326	10	BE359066 D61_38_E0
C 5	39	18.8	444	12	BE918452 OVI_4_A04
C 6	39	18.8	520	13	BI246073 BI246073

7	39	18.8	528	13	BI074353	BI074353 IP1_14_E0
8	39	18.8	529	10	BE358187	BE358187 DG1_26_C0
9	39	18.8	535	12	BE918899	BE918899 FM1_1_E11
10	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
11	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
12	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
13	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
14	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
15	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
16	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
17	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
18	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
19	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
20	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
21	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
22	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
23	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
24	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
25	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
26	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
27	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
28	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
29	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
30	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
31	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
32	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
33	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
34	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
35	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
36	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
37	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
38	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
39	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
40	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
41	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
42	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
43	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
44	39	18.8	537	12	BE918899	BE918899 FM1_1_E11
45	39	18.8	537	12	BE918899	BE918899 FM1_1_E11

ALIGNMENTS

RESULT 1
LOCUS AT170311/c 1282 bp mRNA linear EST 24-JAN-2000
DEFINITION 42 Mycobacterium anaerobic stationary phase library Mycobacterium
smegmatis cDNA, mRNA sequence.
ACCESSION AT170311
VERSION AT170311.1 GI:6742680
KEYWORDS EST.
SOURCE Mycobacterium smegmatis.
ORGANISM Mycobacterium smegmatis.
REFERENCE Murugan-Oel, B., Tay, A. and Dick, T.
TITLE Upregulation of stress response genes and ABC transporters in
anaerobic stationary-phase Mycobacterium smegmatis
JOURNAL Mol. Gen. Genet. 262 (4-5), 677-682 (1999)
MEDLINE 20092472
COMMENT Contact: Murugan-Oel, B.
Mycobacterium Laboratory
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore 117609, Republic of Singapore
Tel: 65 874 3011
Fax: 65 779 1117
Email: mcbomelmb nus.edu.sg
Insert Length: 1282 Std Error: 0.00
Seq primer: T3 Forward: T7 Backward.
Location/Qualifiers
1. 1282

FEATURES
source

```

/organism="Mycobacterium smegmatis"
/db_xref="taxon:1772"
/clone_lib="Mycobacterium anaerobic stationary phase
library"
/lab_host="E. coli XL1-Blue MRF"
/Note="Vector: Lambda ZAP II; Bac111 were disrupted using
a Mini bead beater (Biospec). RNA was isolated using the
RNeasy protocol (Qiagen). Purified RNA was subjected to 2
rounds of digestion with RNase-free DNase I (Promega).
DNase I was heat-inactivated at 75°C for 5 min. and
removed by using RNeasy columns followed by phenol
extraction and ethanol precipitation. The RNA
preparations were confirmed to be free of genomic DNA
contamination by carrying out PCR and RT-PCR using the
Access kit (Promega) and primers specific for the
histone-like protein gene hlp (Lee et al., 1998). cDNA
was synthesized using random hexamer primers (Promega)
and Stratagene's cDNA synthesis kit. cDNA fragments were
ligated into Lambda ZAP II vector and packaged in vitro
using Stratagene reagents."
BASE COUNT      230 a      395 c      405 g      252 t
ORIGIN
Query Match      72.3%  Score 150.4; DB 9; Length 1282;
Best Local Similarity 82.7%  Pred. No. 2e-29;
Matches 172: Conservative 0; Mismatches 36; Indels 0; Gaps 0;

OY 1 TCAGGAGAGAGCGTACACACCTGGCCCGCTGCTGCTTAAGAGTCAACAGAGCTCG 60
    |||||||
DB 236 TCAGGAGAGAGCGTACACACCTGGCCCGCTGCTGCTTAAGAGTCAACAGAGCTCG 177
    |||||||

OY 61 GCGTCGATGTCGCGAGCCATCAGCTGACGCTGACGAGAGAGAGCTGCTGGCCA 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 176 GCGTCGAGCGGCGAAGCGATACACACTGACGCTGACGAGAGAGAGCTGCTCGCA 117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 121 CCATGAATATTTGTCCTGCTTCACAGAGGTCAGACACGATGATCTCCGGCGCG 180
    ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 116 CCATGAGACTGCTGCTGCTGCTGACGAGGTCAGACTGATGATCCCGGTCGCG 57
    ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 181 TCGAGTCCGCGTGAACCGACGACAT 208
    ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 56 TCGAGTCCGCGTGAACCGACGACAT 29
    ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
LOCUS      AO687771/c      877 bp      DNA      linear      GSS 01-JUL-1999
DEFINITION nbx0075101f CUGI Rice BAC Library Oryza sativa genomic clone
ACCESSION  AO687771
VERSION     AO687771.1 GI:5328939
KEYWORDS   GSS.
SOURCE      Oryza sativa.
ORGANISM   Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharitoidae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 877)
AUTHORS   Wing R.A. and Dean R.A.
TITLE     A BAC End Sequencing Framework to Sequence the Rice Genome
JOURNAL   Unpublished (1998)
COMMENT   Contact: Wing RA
          Clemson University Genomics Institute
          Clemson University
          100 Jordan Hall, Clemson, SC 29634, USA
          Tel: 864 656 7288
          Fax: 864 656 4293
          Email: rwing@clemson.edu
          Seq primer: TAATACGACTCACTATAGG
          Class: BAC ends
          High quality sequence stop: 85.
          Location/Qualifiers
FEATURES

```

```

source
1. 877
/organism="Oryza sativa"
/strain="Japonica"
/cu.tivar="Nipponbare"
/db_xref="taxon:4530"
/clone_lib="nbx0075101f"
/clone_lib="CUGI Rice BAC Library"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/Note="Vector: pheloBAC11. Site_1: HindIII. Site_2:
HindIII. Rice is one of two most popular grains in the
world. Half of the world population especially those
inhabiting highly populated areas of the humid tropics
and subtropics, rely on rice as their primary source of
carbohydrate. Monocotyledonous rice is a diploid plant
(2n=24) with a haploid genome equivalent of 431 Mbp
(Arumuganathan and Earle, 1991). The relatively small
genome of rice, three times larger than that of
Arabidopsis, makes it suitable for genomic studies. In
order to facilitate positional cloning, physical mapping
and genome sequencing of rice, we have constructed a BAC
library from Oryza sativa, Nipponbare variety. The
library contains 36,864 clones with an average insert size
of 128.5 kb providing 10.9 haploid genome equivalents. The
deep coverage allows the isolation a particular sequence
with a probability of 99.9%. Two high density filters,
each containing 18,432 clones (doubly spotted), represent
the whole library for colony screening."
BASE COUNT      107 a      370 c      257 g      143 t
ORIGIN
Query Match      20.9%  Score 43.4; DB 17; Length 877;
Best Local Similarity 53.2%  Pred. No. 0.22;
Matches 92: Conservative 0; Mismatches 81; Indels 0; Gaps 0;

OY 29 CGTCGCTCATATAAGTCACAGAAGCTCGGCTCATGTGGCGAGCCCATCAGTC 88
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 236 CGCGCGCGCGGAGGCGACACCGAGCGGTGATACCGTGTGAGTCTCTCCGAC 177
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 89 GTGACGCTGACCGAAGAGACGTGCGCCACCATGATATATCTGCTCCGTTGACGA 148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 176 GTTCCAGGGATGAGGCGGAGCTGAGGCCATGAGAGCCATGAGGCGCCGACGACA 117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

OY 149 GGGTACGACACATGATGATCTCCGGCGCGTGAAGTCCCGTGAACCG 201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 116 GCGCTGCTCCGCGGCGACGACCGCGCGTGGCGCGGAGCTTGCCGAGCGCG 64
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
LOCUS      BI995346      568 bp      mRNA      linear      EST 25-OCT-2001
DEFINITION Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION  BI995346
VERSION     BI995346.1 GI:16430136
KEYWORDS   EST.
SOURCE      Chlamydomonas reinhardtii.
ORGANISM   Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
REFERENCE  1 (bases 1 to 568)
AUTHORS   Grossman A., Chang C.-W., Davies J., Harris E., Hauser C., Lefebvre
          P., McDermost J.P., Shrager J., Sillitov C. and Stern D.
TITLE     Analyses of the Chlamydomonas reinhardtii Genome: A Model,
          Unicellular System for Analyzing Gene Function and Regulation in
          Vascular Plants. Project: 1031
JOURNAL   Unpublished (2001)
COMMENT   Contact: Charles Hauser
          DCMB Box 91000
          Duke University
          Durham, NC 27708-1000
          Tel: 919 613 8159
          Fax: 919 613 8177

```

FEATURES
source

Email: chauser@duke.edu.
Location/Qualifiers

1..568
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized
) lambda zap II"

/note="Vector: Bluescript II SK-; Site_1: EcoRI; Site_2: XhoI; Stress condition II library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH4+ - containing) and shifted to TAP - NO3- (24hrs); H2 production conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda zap clones by superinfection with Exassist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

BASE COUNT 80 a 206 c 185 g 97 t

Query Match 20.3%; Score 42.2; DB 13; Length 568;
Best Local Similarity 50.2%; Pred. No. 0.4;
Matches 104; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTGACCTGCGCGCGCTGCTATAGGTCAACAGAGCTCG 61
DB 251 CAAGGAGAGCGCTGACCTGCGCGCGCTGCTATAGGTCAACAGAGCTCG 310
QY 62 GCTGATGTCGGCGGAGCCCATCATCGTCTGACGCTGACGGAAGAGCTGCGCCAC 121
DB 311 CGTCCGCTACAGGAGGAGCCCGCGCGCTCTCCGCGCGCTCTACCACTGCGCGC 370
QY 122 CATCGAATATCTGTCGCTTGCACAGAGGTCAAGACGATGCTTCGCGGCGCGCT 181
DB 371 CGCGCGCGAGCCCGCGCTTCTCCCGCGCGCGAGCGCGGTTCGCTCCACCTCGCG 430
QY 182 CGAGGTGCGCGGTGCAACCGACGACAT 208
DB 431 CGACAGCGCCGCTGGCGGCTTGGCAT 457

RESULT 4
BE359066 326 bp mRNA linear EST 20-JUN-2000
LOCUS DGI.38.E06.g1_A002 Dark grown 1 (DGI) Sorghum bicolor cDNA, mRNA
DEFINITION sequence.
ACCESSION BE359066
VERSION BE359066.1 GI:9300623
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 326)
REFERENCE Cordomier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
AUTHORS L.H.

TITLE An EST database from Sorghum: dark-grown seedlings
JOURNAL Unpublished (2000)
COMMENT Contact: Cordomier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: Polymix
High quality sequence start: 5
High quality sequence stop: 324
POLYA=Yes.

FEATURES
source

Location/Qualifiers
1..326
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DGI)"

/note="Organ: 5-day-old dark-grown seedlings; Vector: lambda zap; Site_1: XhoI; Site_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda zap II. Clones to be sequenced were prepared by mass excision."
BASE COUNT 72 a 93 c 118 g 43 t

Query Match 18.8%; Score 39; DB 10; Length 326;
Best Local Similarity 50.3%; Pred. No. 2.5;
Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACGACCTGCGCGCGCTGCTATAGGTCAACAGAGCTCG 61
DB 117 CAAGTGGTGTCTACAGCTGCTGAGAGCGGAGCTGTTACTTACAGAACTTCAA 176
QY 62 GCTGATGTCGGCGGAGCCCATCATCGTCTGACGCTGACGGAAGAGCTGCGCCAC 121
DB 177 GATCAGCTGAAGGTGCGAGGCGCGCGCGCGAGCGCGGTAGTATGCGCGCGCT 236
QY 122 CATCGAATATCTGTCGCTTGCACAGAGGTGACAGCAGATGCTTCGCGGCGCGCT 181
DB 237 GCTCAATGTCATATGAGCTTCAGCAAGCCAGACAGAGTGCACCGCGGACGTAT 296
QY 182 CGAGTGGCGG 192
DB 297 CAAGGAGACG 307

RESULT 5
BE918452 444 bp mRNA linear EST 29-SEP-2000
LOCUS OV1_4.A04.g1_A002 Ovary 1 (OV1) Sorghum bicolor cDNA, mRNA
DEFINITION sequence.
ACCESSION BE918452
VERSION BE918452.1 GI:10421570
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 444)
REFERENCE Cordomier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
AUTHORS L.H.

TITLE An EST database from Sorghum: ovaries of varying immature stages
JOURNAL Unpublished (2000)
COMMENT Contact: Cordomier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: Polymix
High quality sequence start: 90
High quality sequence stop: 427
POLYA=No.
Location/Qualifiers

FEATURES

```

source
1..444
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Ovary 1 (OV1)"
/note="Organ: Mix of ovaries of varying immature stages
from 8-week-old plants; Vector: pBluescript II from Lambda
Zap II; Site_1: XhoI; Site_2: EcoRI; The library was made
from poly-A RNA in the cloning vector Lambda Zap II.
Clones to be sequenced were prepared by mass excision."
BASE COUNT
102 a 131 c 142 g 69 t
ORIGIN

```

	Query Match	18.8%	Score 39	DB 12	Length 444
	Best Local Similarity	50.3%	Pred. No.	2.6	
	Matches	96	Conservative	0	Mismatches 95
					Indels 0
					Gaps 0
Oy	2	CAMGAGACGCGTACGACCTGGCCCGCGTGGCTGATATAGGTCAACAAGACTCGG	61		
Db	162	CAMGGTGGTCTGCTACTGCGTGTGAGAGGCGACTGTGATCTTACAAAGACTTCAA	221		
Oy	62	GCAGCATGTGGGGGACCCCATCAGTGTGTGACGCTGACCGAAGAACACTCGTGGCCAC	121		
Db	222	GATCAGCGCTAAGGTGGCCCAAGCGGGCGGAGGCGCAGGGGTGAGGCCGCCGCCCT	281		
Oy	122	CATCGAATATCTGTGCCGCTTGACAGAGGTCACAGACCATGATCTGTCGGCGGGCGT	181		
Db	282	CGTCACTGGTCCATGAGATGTCGACAAGGCCAACAGACAGAGTGCCGACCCGACGTCA	341		
Oy	182	CGAGGTGCCCG	192		
Db	342	CAMGAGACCG	352		

FEATURES	source
RESULT 6	
LOCUS	B1246073
DEFINITION	IP1.68.D09_b1.A002 Immature pannicle 1 (IP1) Sorghum bicolor cDNA,
ACCESSION	B1246073
VERSION	B1246073.1
KEYWORDS	GI:14824091
SOURCE	EST.
ORGANISM	sorghum.
REFERENCE	Sorghum bicolor
AUTHORS	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae: PACC clade: Panicoideae: Andropogoneae; Sorghum.
TITLE	1 (bases 1 to 520)
JOURNAL	Klein,R.R., Cordonnier-Pratt,M.,M., Gingle,A., Sudman,M. and Pratt, L.H.
COMMENT	An EST database from Sorghum: developing preanthesis pannicles Unpublished (2001) Contact: Cordonnier-Pratt MM Laboratory for Genomics and Bioinformatics The University of Georgia, Department of Plant Biology Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 583 0210 Email: mmp@prattuga.edu Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for high quality sequence is 20. Three-prime sequences, which are obtained with Polymix or T7 sequencing primer, are presented as the reverse complement. Seq primer: JEN REV High quality sequence stop: 400 PolyA-No.
FEATURES	Location/Qualifiers
source	1..520

```
FEATURES
source
    Location/Qualifiers
        1..520
            /organism="Sorghum bicolor"
            /cultivar="BTx623"
            /db_xref="taxon:4558"
            /clone_1fb="Immature panicle1 (1p1)"
            /note="Organ: Developing preleishe panicles; Vector:
            bluescript II SK(-) from Lambda zap II; Site1: XhoI;
```

Query Match	18.8%	Score 39;	DB 13;	Length 520;
Best Local Similarity	50.3%;	Pred. No. 2.7;		
Matches	96;	Conservative	0;	Mismatches 95; Indels 0; Gaps 0;
BASE COUNT	113 a	153 c	178 g	76 t
ORIGIN	Site 2: ECORI. The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision.*			

QY 2 CAAAGGAAACCGCTAGACGCTTCCGCCCGGTGGGCGCTATTAAGGTCAACAAAGAACTCGG 61
 Db 300 CAAAGGTGTGTGCTTACAGCGTGTGTGACGCGCAACTGTTGACTTCTTACAAAGAACTTCAA 359
 QY 62 GCTGCATGTCGGCGAAGCCCATCAGTCGTCACGCTGACCGAAGAAAGACGTGTGGCCAC 121
 Db 360 GATCAGCGCTAAAGTGTGGCCAAAGCGGGCGGGGAGAGGGGTGAGAGCCCGGCCCT 419
 QY 122 CATCGAATATCTGTGTCCGCTTGACAGAGGTCAGACCAACAGATATCGTTCCGGGCGCGGT 181
 Db 420 CGTCACTGTGTCATGAGATTCGACAAAGGCCAAAGCAAGGTCCCGACCCGGAGCTAT 479
 QY 182 CGAGGTGCCG 192
 Db 480 CAAAGAGACCG 490

RESULT 7	
B1074353	
LOCUS	B1074353
DEFINITION	528 bp mRNA linear EST 20-JUN-2001
VERSION	IP1.14.E03.g1.A002 Immature panicle 1 (IP1) Sorghum bicolor cDNA,
ACCESSION	mRNA sequence.
KEYWORDS	B1074353 B1074353 GI:14513010
SOURCE	EST.
ORGANISM	sorghum.
	Sorghum bicolor

REFERENCE	TITLE
1 (bases 1 to 528)	An EST database from Sorghum: developing preanthesis panicles
Klein, R.R., Cordonnier-Pratt, M.-M., Gingle, A., Sudman, M. and Pratt, L.H.	Unpublished (2001)
	Contact: Cordonnier-Pratt MM

The University of Georgia, Department of Plant Biology,
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmp@attpuga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for high quality sequence is 20. Three-prime sequences, which are obtained with PolyMix or T7 sequencing primer, are presented as the reverse complement.

Seq primer: T7

High quality sequence start: 13

High quality sequence stop: 528

PolyA-yes.

FEATURES	SOURCE
Location/Qualifiers	1..528
/organism="Sorghum bicolor"	
/cultivar="PR623"	
/db_xref="taxon:4558"	
/clone_l1b="Immature panicle 1 (P1)"	
/note="Organ: Developing preanthesis panicles; Vector: pBluescript II SK(-) from lambda Zap II; Site.1: XhoI; Site.2: EcoRI; The library was made from poly(A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."	
BASE COUNT	103 a 147 c 164 g 114 t
BRIGIN	

Query Match	18.8%	Score 39;	DB 13;	Length 528;
Best Local Similarity	50.3%;	Pred. No. 2.7;		
Matches 96;	Conservative 0;	Mismatches 95;	Indels 0;	Gaps 0

OY 2 CAAGGAGAAGCCCTACCACTTGCCCCCGTCGGTGCCTATAAGGTCAACAAAGAAGCTCGG 61
 ||||| - - - ||| - ||| - | - | ||||| ||
Db 29 CAAAGTGGTGTCTACAGCGTGGTGAACGGCAGCTGTTGACTTCTACAAGAAGCTTCAA 88

Dy 62 GCTGCATGTCGGCGAGGCCCATACGTCTGACCCTGACCGAAGAAGACGTGTGGCCAC 12

Db 89 GATTCACGCTGAAGGTGGCCAAAGCGGCGCGGCGGAGGCGAGGCTGAGGCCGCGCGT 14

QY 122 CATCGAATATCTGTCCTCCGTCACGAGGGTCACACCACCATGATCTTCCGGCGGGGT 18
 149 CGTCAACTGGTCCATGGAGTTTCGACAAAGCCAAACGACCAAGTGCCTGACCCGAGCGTCAAT 20

QY	182	CGAGGTGCCCG	192
Db	209	CAAGGAGACCG	219

RESULT 8
BE358187
LOCUS
nnnnnnnnnn
555

DEFINITION	ACCSSION	VERSION
DG1_26_C02.b2_A002 Dark Grown 1 (DG1) Sorghum bicolor cDNA, mRNA sequence.	BE358187	1
	BE358187	1
	BE358187	1

KEYWORDS	EST.
SOURCE	sorghum.
ORGANISM	Sorghum bicolor

Spermatophyta: Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.

TITLE An EST database from Sorghum: dark-grown seedlings
JOURNAL Unpublished (2000)
AUTHOR L.H.
CONTACT Contact: Abdelhak Ben Drah, INRA, UR1013, 17130 St-Paul, France

Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
e-mail: 705.542.1050

Fax: 706 583 0210
Email: mmprat@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions between standard mutations.

Seq primer: JEN REV
High quality sequence stop: 388

FEATURES	Location/Qualifiers
source	1. .529
	/organism="Sorghum bicolor"

```

/cione_1lb="Dark Grown 1 (DG1)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
Lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was

```

BASE COUNT	ORIGIN	Clones to be sequenced were prepared by mass excision.
116 a	156 c	182 g
		75 t

Query Match	18.8%	Score 39;	DB 10;	length 529;
Best Local Similarity	50.3%;	Pred. No. 2.7;		
Matches	96;	Conservative	0;	Mismatches 95;
			Indels	0;
			Gaps	0;

2 CAAGGAGAGCCCTACGACCTGGCCCGCGTCGGTATTAAGGTCACAGAGAGCTCGG 61
||||| | ||| |||| | || | | || ||||| ||
312 CAAGTGCTGCTACACAGCTGGTGGAGCGCGACCTGGTTGACTTCTACAGAACTTCAA 371

62 GCTGATGTCCGGAGCCCCATCACGTCGTCCAGCGCTGACCCGAGAAGACGTCGTGGCCAC 121
 121
 372 GATCACGCTGAAGTGTGGCCAAAGCGCGCGCGGAGAGGGCGAGGAGGTGAGCGCCGCGCGCT 431
 431

122 CATTGATATCTGGTCCGCTTGACAGAGGGTTCAGACACAGATGATCGTTCCGGGGCGGCT 181
 432 CGTCAACTGATCCATGGAGTTGACAAAGGCCAACGACCAGGTGCCCGACCCGGAGCTCAT 491

x7	102	CAGGAGACCG	192
D6	492	CAGGAGACCG	502

LOCUS	BE918899	535 bp	mpna	1 hour	ecm 20-cm-2000
RESULT 9	BE918899				

DEFINITION	ACCESSION	VERSION
FM1-1_E11.B1_A003 Floral-Induced Meristem 1 (FM1) Sorghum protopinquinum cDNA, mRNA sequence.	BE918899	BE918899.1 GI:10422437

REINFORCED	EST.
SOURCE	Sorghum propinquum.
ORGANISM	Sorghum propinquum
	Elkarvota. Viridilantae: Streptococcus. Nubantubus

REFERENCE

GORDONIER-PRATT M.-M., CIOLO A., SUDAN N., MORILLAS J.-C. and BOUTIN P. 1987, Les plantes vasculaires de la région méditerranéenne occidentale.

TITLE	An EST database from Sorghum: floral-induced meristems
JOURNAL	Unpublished (2000)
COMMENT	Contact: Cordonnier-Pratt MW
	, L.H.

Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860

Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for high-quality

15 20.
Seq primer: JEN REV
High quality sequence stop: 381
POLYA-No.

FEATURES	Location/Qualifiers
SOURCE	1. 535
	/organism="Sorghum propinquum"
	/db xref="Rayon:132711"

```

/cone_1lb="Floral-Induced Meristem 1 (FM1)"
/note="Organ: Floral-Induced meristems; Vector:
pluInscript II from Lambda zap II; Site_1: XhoI; Site_2:
FcoRI; mature plants were raised from seedlings

```

15 days with 16 hr darkness and 8 hr light (flowering is induced by short-day conditions); 16 days after being returned to the greenhouse under natural long days during late April/early May. Rootstocks were maintained under

library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."

ORIGIN	Query Match	Score 39;	DB 12;	Length 535;
Best local similarity	18.8%	50.2%	52.2%	

Matches	96;	Conservative	0;	Mismatches	95;	Indels	0;	Gaps	0;
2	CAAGGACAGCCCTACGACCTGGCCCCGCGTGGTGGTATAAAGTCAACAGAAAGCTCGG	61							

308 CAAGGTGGTGTGTACAGCGTGGTGCAGGGCCAGCTGGTTGACTTCTACAAAGCACTTCAA 367

368 GATCACGCTGAAGGTGGCCAAAGCGCGCGCGCGGAGGGCGAAGGTGAGGCCCGGCCCGT 427

QY 122 CATCAATATCTGTCCTTGCACAGAGGTGCAGACCAGTATCGTCCGGCGCGT 181
 DB 428 CGTCAACTGTCATGAGCTTCGACAAAGCCACAGCAGCAGTCCCGACCGGACGTAT 487
 QY 182 CGAGTCCCG 192
 DB 488 CAAGGAGACCG 498

RESULT 10
 LOCUS BG412695 537 bp mRNA linear EST 13-MAR-2001
 DEFINITION OV2_31_F06.b1_A002 Ovary 2 (OV2) Sorghum bicolor cDNA, mRNA
 sequence.
 ACCESSION BG412695
 VERSION BG412695.1 GI:13318248
 KEYWORDS EST.
 SOURCE Sorghum.
 ORGANISM Sorghum bicolor

REFERENCE Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta;
 Spermatophyta: Magnoliophyta: Liliopsida: Poales; Poaceae; PACC
 clade: Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 537)
 Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
 ,L.H.

TITLE An EST database from Sorghum: ovaries of varying immature stages
 JOURNAL Unpublished (2000)
 COMMENT Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmp@prattuga.edu
 Seq primer: JEN REV
 High quality sequence stop: 485
 POLYA-No.

FEATURES

source Location/Qualifiers
 1..537

/organism="Sorghum bicolor"
 /db_xref="taxon:4558"
 /clone_lib="Ovary 2 (OV2)"

/note="Organ: Mix of ovaries of varying immature stages
 from 8-week-old plants. Vector: pBluescript II from Lambda
 Zap II; Site_1: XhoI; Site_2: EcoRI; The library was made
 from poly-A RNA in the cloning vector lambda Zap II.
 Clones to be sequenced were prepared by mass excision."
 BASE COUNT 117 a 159 c 183 g 78 t

Query Match 18.8%; Score 39; DB 12; Length 537;
 Best Local Similarity 50.3%; Pred. No. 2.7;

Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 2 CAAGAGAAGCCCTACGACCTGCGCCGCGTGTATTAAGGTCAACAGAGCTCG 61
 DB 306 CAAGGTGCTGTCATGACGCTGTGACGCGGAGCTGTTGACTTACAGAACTTCA 365
 QY 62 GCTGATGTCGGCGAGCCCATCATCGTCTGACGCTGACCGAAGAGAGCTGTGCCAC 121
 DB 366 GATCAGCGCTGAAGGTGCGCAAGCGCGCGGAGGCGAGGGTGAAGCGCGCGCT 425
 QY 122 CATCGAATATCTGTCCTTGCACAGAGGTGCAGACCAGTATCGTCCGGCGCGT 181
 DB 426 CGTCAACTGTCATGAGCTTCGACAAAGCCACAGCAGTCCCGACCGGACGTAT 485
 QY 182 CGAGTCCCG 192
 DB 488 CAAGGAGACCG 496

RESULT 11

BG649762
 LOCUS BG649762 537 bp mRNA linear EST 24-APR-2001
 DEFINITION RH12_83_C10.b1_A003 Rhizome2 (RH12) Sorghum prolinguam cDNA, mRNA
 sequence.
 ACCESSION BG649762
 VERSION BG649762.1 GI:13784874
 KEYWORDS EST.
 SOURCE Sorghum prolinguam.
 ORGANISM Sorghum prolinguam

REFERENCE Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta;
 Spermatophyta: Magnoliophyta: Liliopsida: Poales; Poaceae; PACC
 clade: Panicoideae; Andropogoneae; Sorghum.
 1 (bases 1 to 537)
 Cordonnier-Pratt,M.-M., Gingle,A., Paterson,A., Sudman,M. and Pratt
 ,L.H.

TITLE An EST database from Sorghum: Sorghum prolinguam rhizomes
 JOURNAL Unpublished (2000)
 COMMENT Contact: Cordonnier-Pratt MM
 Laboratory for Genomics and Bioinformatics
 The University of Georgia, Department of Plant Biology
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
 Tel: 706 542 1860
 Fax: 706 583 0210
 Email: mmp@prattuga.edu
 Seq primer: JEN REV
 High quality sequence stop: 396
 POLYA-No.

FEATURES

source Location/Qualifiers
 1..537

/organism="Sorghum prolinguam"
 /db_xref="taxon:132711"
 /clone_lib="Rhizome2 (RH12)"

/note="Organ: Rhizomes; Vector: pBluescript II from Lambda
 Zap II; Site_1: XhoI; Site_2: EcoRI; The library was made
 from poly-A RNA in the cloning vector lambda Zap II.
 Clones to be sequenced were prepared by mass excision."
 BASE COUNT 119 a 158 c 181 g 79 t

Query Match 18.8%; Score 39; DB 12; Length 537;
 Best Local Similarity 50.3%; Pred. No. 2.7;

Matches 96; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 2 CAAGAGAAGCCCTACGACCTGCGCCGCGTGTATTAAGGTCAACAGAGCTCG 61
 DB 296 CAAGGTGCTGTCATGACGCTGTGACGCGGAGCTGTTGACTTACAGAACTTCA 355
 QY 62 GCTGATGTCGGCGAGCCCATCATCGTCTGACGCTGACCGAAGAGAGCTGTGCCAC 121
 DB 356 GATCAGCGCTGAAGGTGCGCAAGCGCGCGGAGGCGAGGGTGAAGCGCGCGCT 415
 QY 122 CATCGAATATCTGTCCTTGCACAGAGGTGCAGACCAGTATCGTCCGGCGCGT 181
 DB 416 CGTCAACTGTCATGAGCTTCGACAAAGCCACAGCAGTCCCGACCGGACGTAT 475
 QY 182 CGAGTCCCG 192
 DB 476 CAAGGAGACCG 486

RESULT 12
 LOCUS BE917902 550 bp mRNA linear EST 29-SEP-2000
 DEFINITION OVL_7_E07.b1_A002 Ovary 1 (OVL) Sorghum bicolor cDNA, mRNA
 sequence.

ACCESSION BE917902
 VERSION BE917902.1 GI:10420470
 KEYWORDS EST.
 SOURCE Sorghum.
 ORGANISM Sorghum bicolor

REFERENCE Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta;
 Spermatophyta: Magnoliophyta: Liliopsida: Poales; Poaceae; PACC
 clade: Panicoideae; Andropogoneae; Sorghum.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 ; Search time 21.1607 Seconds
(Without alignments)
3487.380 Million cell updates/sec

Title: US-09-697-123b-22

Perfect score: 208
Sequence: 1 tcaaggagaagcgctacgac.....ccgltggaacgcagacat 208

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	42.3	5096	10 US-09-984-711-5	Sequence 5, Appli
2	88	42.3	5099	9 US-10-075-460-5	Sequence 5, Appli
3	88	42.3	5099	10 US-09-887-052-1	Sequence 1, Appli
4	88	42.3	5099	10 US-09-887-052-3	Sequence 3, Appli
5	88	42.3	5099	10 US-09-887-052-5	Sequence 3, Appli
6	35.6	17.1	3756	10 US-09-841-132-425	Sequence 425, App
7	34	16.3	252	10 US-09-878-574-8901	Sequence 8901, Ap
8	34	16.3	680	10 US-09-841-132-267	Sequence 267, App
9	32.4	15.6	1437	10 US-09-815-242-7738	Sequence 7738, Ap
10	32	15.4	711	10 US-09-815-242-7723	Sequence 7723, Ap
11	31.4	15.1	11220	10 US-09-815-242-7696	Sequence 7696, Ap
12	31.4	15.1	11220	10 US-09-861-289-32	Sequence 32, Appli
13	31.4	15.1	36778	10 US-09-861-289-5	Sequence 5, Appli
14	31.2	15.0	278	10 US-09-294-093B-1392	Sequence 1392, Ap
15	31.2	15.0	704	12 US-10-062-254-109	Sequence 109, App
16	31.2	15.0	752	12 US-10-062-254-111	Sequence 111, App
17	31.2	15.0	936	10 US-09-815-242-7776	Sequence 7776, Ap
18	31.2	15.0	8673	10 US-09-815-242-7474	Sequence 7474, Ap
19	31	14.9	1287	12 US-10-029-715-3	Sequence 3, Appli

20	31	14.9	3084	10 US-09-881-752A-79	Sequence 79, Appli
21	31	14.9	8673	10 US-09-815-242-7313	Sequence 7313, Ap
22	30.8	14.8	337	10 US-09-960-352-10223	Sequence 10223, A
23	30.8	14.8	1293	10 US-09-815-242-7978	Sequence 7978, Ap
24	30.8	14.8	1296	10 US-09-815-242-4090	Sequence 4090, Ap
25	30.8	14.8	1440	10 US-09-796-858-7	Sequence 7, Appli
26	30.8	14.8	1534	10 US-09-765-831A-20	Sequence 20, Appli
27	30.6	14.7	264	10 US-09-923-876-4989	Sequence 4989, Ap
28	30.6	14.7	953	10 US-09-452-239-45	Sequence 45, Appli
29	30.6	14.7	1266	10 US-09-815-242-7920	Sequence 7920, Ap
30	30.2	14.5	585	10 US-09-870-162A-10	Sequence 10, Appli
31	30.2	14.5	1298	10 US-09-825-300-682	Sequence 682, App
32	30.2	14.5	2109	10 US-09-815-242-7761	Sequence 7761, Ap
33	30.2	14.5	6252	10 US-09-964-824A-23	Sequence 23, Appli
34	30.2	14.5	14272	10 US-09-870-162A-23	Sequence 313, App
35	30	14.4	1455	10 US-09-887-576-782	Sequence 782, App
36	30	14.4	2329	10 US-09-816-828-9	Sequence 9, Appli
37	30	14.4	4863	10 US-09-815-242-4071	Sequence 4071, Ap
38	29.8	14.3	592	10 US-09-894-467-6	Sequence 6, Appli
39	29.8	14.3	1305	10 US-09-815-242-7730	Sequence 7730, Ap
40	29.8	14.3	2430	10 US-09-861-289-23	Sequence 23, Appli
41	29.8	14.3	13613	10 US-09-815-242-4001	Sequence 4001, Ap
42	29.6	14.2	897	10 US-09-964-824A-553	Sequence 553, App
43	29.6	14.2	914	10 US-09-880-107-2357	Sequence 2357, Ap
44	29.6	14.2	914	10 US-09-981-353-126	Sequence 126, App
45	29.6	14.2	1115	9 US-09-981-353-126	

ALIGNMENTS

RESULT 1
US-09-984-711-5
; Sequence 5, Application US/09984711
; Patent No. US20020119549A1
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: BATHE, Brigitte
; APPLICANT: STEPHAN, Hans
; APPLICANT: KREUTZER, Caroline
; APPLICANT: HERMANN, Thomas
; APPLICANT: PEEFERLE, Walter
; APPLICANT: BINDER, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE
; FILE REFERENCE: 204209US0
; CURRENT APPLICATION NUMBER: US/09/984,711
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: DE10108230.9
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 5096
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
; OTHER INFORMATION:
US-09-984-711-5

Query Match 42.3%; Score 88; DB 10; Length 5096;
Best Local Similarity 68.5%; Pred. No. 3.5e-17;
Matches 137; Conservative 0; Mismatches 60; Indels 3; Gaps 1;

Qy 9 AAGGCGACACCGTGGCGGCGTGGCTATTAAGCTCAACGAAGCTGGCTGCAT 68
Db 1578 AAGCGCTACACCGTGGCGGCGTGGCTGTGTACAAATCAACGCAAGCTGGCT--T 1634

Qy 69 GTGGCGAGCCATCATCGTGTGAGCGTGAACGAGAGAGAGAGTGTGGCCACATCGAA 128
Db 1635 GTGGCGAGCCATCATCGTGTGTGAGTCTTACTGAGAGAGAGATCGCAACATCGAG 1694

Db 1635 GGTGGGACACAGATGGTTGATGACTCTTACTGAAGAGAGATGCAACCAACGCTTCAG 1694
QY 129 TATCTGTCCTGCTTGCACGAGGTCAGACCAGATGATGTTCCGGCGGCTCGAGGTG 188
Db 1695 TACCTGTGCTGCTGCACGAGGTGAGCGGCTCATGATGATCTTCCAAATGGTGAAGAGATC 1754
QY 189 CCGGTGGAACCGACGACAT 208
Db 1755 CCACTGCAGACCATGACAT 1774

RESULT 5

US-09-887-052-5
; Sequence 5, Application US/09887052
; Patent No. US20020119537A1
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: BATHE, Brigitte
; APPLICANT: HERMANN, Thomas
; APPLICANT: PEEFERLE, Walter
; APPLICANT: BINDER, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE
; FILE REFERENCE: 204212050X
; CURRENT APPLICATION NUMBER: US/09/887,052
; PRIOR APPLICATION NUMBER: 2001-06-25
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
US-09-887-052-5

Query Match 42.3%; Score 88; DB 10; Length 5099;
Best Local Similarity 68.5%; Pred. No. 3.5e-17;
Matches 137; Conservative 0; Mismatches 60; Indels 3; Gaps 1;

QY 9 AAGCCCTACGACCTGGCCCGGCTGCTATAGGTCAACAAGAGCTCGGGCTCAT 68
Db 1578 AAGCCCTACGACCTGGCTGCTGCTTGAACAAGATCAACCCAGCTCGGGCT---T 1634
QY 69 GTGGGCGAGCCCATCAGCTCTGTCAGCGTGAACGAGAGAGCTGTGGCCACCATCGAA 128
Db 1635 GTGGGCGAGCCCATCAGCTCTGTCAGCGTGAACGAGAGAGCTGTGGCCACCATCGAG 1694
QY 129 TATCTGTCCTGCTTGCACGAGGTCAGACCAGATGATGTTCCGGCGGCTCGAGGTG 188
Db 1695 TACCTGTGCTGCTGCACGAGGTGAGCGGCTCATGATGATCTTCCAAATGGTGAAGAGATC 1754
QY 189 CCGGTGGAACCGACGACAT 208
Db 1755 CCACTGCAGACCATGACAT 1774

RESULT 6

US-09-841-132-425
; Sequence 425, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121,469C8
; CURRENT APPLICATION NUMBER: US/09/841,132
; CURRENT FILING DATE: 2001-04-23

; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 425
; LENGTH: 3756
; TYPE: DNA
; ORGANISM: Chlamydia trachomatis serovar D
US-09-841-132-425

Query Match 17.1%; Score 35.6; DB 10; Length 3756;
Best Local Similarity 56.5%; Pred. No. 0.095;
Matches 87; Conservative 0; Mismatches 64; Indels 3; Gaps 1;

QY 1 TCAGAGGAGAGCGCTACGACCTGGCCCGGCTGCTGCTATAGGTCAACAAGAGCTCG 60
Db 983 TCGATCTTAACGCTATATATTAGCCCGGCTTGACGTTATTAATTAATAAATTAG 1042
QY 61 GCGTGC---ATGTGGCGAGCCCATCAGCTGTCGACGCTTACCGAAGACGCTGTG 117
Db 1043 GATTCCTATTACGACGAGAAACATATCTCAAGTCACTTGAAGAAAGAGATGTATCG 1102
QY 118 CCACATCGAATATCTGTCGCTGCTGACAGAGG 151
Db 1103 GCGCGTTGAATATTTGATTCGTTGCAATGG 1136

RESULT 7

US-09-878-574-8901
; Sequence 8901, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR APPLICATION NUMBER: 2001-12-21
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 8901
; LENGTH: 252
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701101828H1
US-09-878-574-8901

Query Match 16.3%; Score 34; DB 10; Length 252;
Best Local Similarity 50.0%; Pred. No. 0.16;
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 17 GACCTGGCCCGGCTGCTGCTATAGGTCAACAAGAGCTGGGCTGATGCTGGCGA 76
Db 82 GCGCGTGGCCCGCATGAGCGACCCGACGCTCATCAAGCCATACGCGGCGCTCACCAT 141
QY 77 GCCCATCAGCTGTCGACGCTGACCGAAGAGAGCTGTGGCCACCATCAATATCTGT 136
Db 142 CCCGCTATGCGCCAGCGCCGACATCGGCTGTCGAGAGCCCAATCTCGAAGGCCCT 201
QY 137 CCGCTGACGAGGGTCAAGCACGATGATCTTCGGGCGGCGGTGAGG 186
Db 202 CCGCATGACATCTGACGACGCGCGGAGGTCCTACCTTCGCGAGGAGG 251

RESULT 8

US-09-841-132-267
; Sequence 267, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
;; FILE REFERENCE: 210121.469c8
;; CURRENT APPLICATION NUMBER: US/09/841,132
;; NUMBER OF SEQ ID NOS: 599
;; SOFTWARE: FastSeq for Windows Version 3.0/4.0
;; SEQ ID NO 267
;; LENGTH: 680
;; TYPE: DNA
;; ORGANISM: Chlamydia
US-09-841-132-267

Query Match 16.3%; Score 34; DB 10; Length 680;
Best Local Similarity 55.8%; Pred. No. 0.2;
Matches 86; Conservative 0; Mismatches 65; Indels 3; Gaps 1;

OY 1 TCAAGAGAGCCGTACGACCTGGCCGCTGCTGCTATAGTCAACAAGAGCTCG 60
DB 442 TCGATCGTAAACCTATTAATTAGCGCGGTGACGTTATAAATTAATAAATTAG 501
OY 61 GGTGCG--ATGTCGGCGAGCCCATACGTCGTCGACGCTGACCGAAGAGCTGCTGG 117
DB 502 GCTTCCCATTTAGACGACGAACCTATTCACAGTACTTTGAGAAAAAGAGATGTTATCG 561
OY 118 CCACCATCGAATATCTGTCGCTGCTGACGAGGG 151
DB 562 GCCGCTTGAATAATTTGATTGCTTTCGAATGG 595

RESULT 9
US-09-815-242-7738
; Sequence 7738, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7738
; LENGTH: 1437
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1437)
US-09-815-242-7738

Query Match 15.6%; Score 32.4; DB 10; Length 1437;
Best Local Similarity 52.2%; Pred. No. 0.69;
Matches 72; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

OY 46 TCACACAGAGCTGGCGGTGATGTCGGCGAGCCCATCAGTCGTCGACGCTGACCCGAG 105
DB 809 TCGACAGACTGATCTGTCGCTGCGCCGCTGCGCCGAGCACCGACGCTGCGCCCGC 868
OY 106 AAGACGTCGTGCGCCACATCGAATATCTGTCGCTTGACAGAGGTGACGACGATGA 165
DB 869 ACAGCGGCGTACACCTGGACAGAGCGTGGCTTATCTAGCTGACGACACACTGCAAGACA 928
OY 166 TCGTTCGGCGCGGCTCG 183
DB 929 GCGTTCGGCGGCTCTTCG 946

RESULT 10
US-09-815-242-7723
; Sequence 7723, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7723
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(711)
US-09-815-242-7723

Query Match 15.4%; Score 32; DB 10; Length 711;
Best Local Similarity 55.4%; Pred. No. 0.78;
Matches 62; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

OY 92 GAGCGTACCCGAGAGAGCTGTCGGCCACCATCATATATCTGTCGCTGACGAGG 151
DB 171 GATGCTGCGCGGAGAGAGCGCTGTCATCTGCCCGCAGTGCGCGAGTTGACGG 230
OY 152 TCAGACACGATGATGCTTCCGGGCGGCGAGGTGCGCGGTGGAACGAC 203
DB 231 TCGATCTGATGCTCACCGCGCGCACCGATATGAGACGAGTACTCGC 282


```
RESULT 11
US-09-815-242-7696
; Sequence 7696, Application US/09815242
; Patent No. US2002006169A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7696
; LENGTH: 1164
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1164)
US-09-815-242-7696

Query Match      15.1%; Score 31.4; DB 10; Length 1164;
Best Local Similarity 48.1%; Pred. No. 1.3;
Matches 89; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 23 GGGCCGCTGCGTGTATAGGTCAACAAGACTCGGGCGCATGTGCGCGAGCCCAT 82
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
DB 558 GGGGATCGTCGGCGGCTCCAGAGTTTCCACCAGCTCGACGCTGGAACAGCTGGCGG 617
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
QY 83 CACGTCGTGACGCTGACCGAAGAGCTGCGGCACCATCGAATATCTGTGCGGCTT 142
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
DB 618 CATCTCGACCAACGATGTGCGGGGCGGATCGCCCAACCTTCTCGCGCGCGCGG 677
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
QY 143 GCACGAGGCTACAGACCATATATGTTCCGGCGCGCTCGAGTGCCTGGAACCGA 202
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
DB 678 ACACAGCTGCGCAAGTACTGTACGAGCGGACCTGTGCGAGACCGGACGCGCATTCG 737
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||

QY 203 CGACA 207
    ||| ||
DB 738 CGCCA 742

RESULT 12
US-09-861-289-32/c
; Sequence 32, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
```

```
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 11220
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-861-289-32
```

```
Query Match      15.1%; Score 31.4; DB 10; Length 11220;
Best Local Similarity 48.1%; Pred. No. 2.1;
Matches 89; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
```

```
QY 12 CGCTACGACCTGCGCCCGCTGCTGCTATTAAGTCACACAGAGCTCGGCTGCATGTC 71
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
DB 5535 CTCGACGACGCTCAGCTGCGCGGCTGCTGAGTGGGCTCGGAGCGCGCGGATGAC 5476
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
QY 72 GCGGAGCCCATCAGCTGCTGACGCTGACCGAAGAACGTCGTGCGCCACCATCGAATAT 131
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
DB 5475 GCGCTGCTGCGAGGCGCCGCTTCCGGCGGCTCAGGCGCTGCTCGCGCGCTCGTTGAC 5416
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
QY 132 CTGTCGCTTTCGACGAGGAGTACAGACCATGATGCTTCCGGGCGGCTCGAGTGGCG 191
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
DB 5415 CCGGCTGCGCGCGACGACCGCCACGACGACGATGCTCGTTCGGCGGCGGCTCGACAGCGG 5356
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
QY 192 CTGCA 196
    ||| ||
DB 5355 CTGCA 5351
```

```
RESULT 13
US-09-861-289-5/c
; Sequence 5, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/861,289
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 36778
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-861-289-5
```

```
Query Match      15.1%; Score 31.4; DB 10; Length 36778;
Best Local Similarity 48.1%; Pred. No. 2.6;
Matches 89; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
```

```
QY 12 CGCTACGACCTGCGCCCGCTGCTGCTATTAAGTCACACAGAGCTCGGCTGCATGTC 71
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
DB 21222 CTCGACGACGCTCAGCTGCGCGGCTGCTGAGTCGCGCGCTCGCGCGCGCGGATGAC 21163
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
QY 72 GCGGAGCCCATCAGCTGCTGACGCTGACCGAAGAACGTCGTGCGCCACCATCGAATAT 131
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
DB 21162 GCGCTGCTGCGAGGCGCCGCTTCCGGGCGGCTCAGGCGCTTCTCGCGCGCTGAGTTGAC 21103
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
QY 132 CTGTCGCTTTCGACGAGGAGTACAGACCATGATGCTTCCGGGCGGCTCGAGTGGCGG 191
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
DB 21102 CCGGCTGCGCGCGACGACCGCCACGACGATGCTCGTTCGGGCGGCGGCTCGACAGCGG 21043
    ||| ||||| ||| ||||| ||| ||||| ||| ||| |||
```

PRIOR FILING DATE: 2000-07-28

07 b4 TGCATGTGGGAGCCCATCACTGTCGACGCTGACCGAGAAGACG 111
 | | | | | | | | | |
 Db 335 TCGTCTACCAGCTCAACATCGCCCCCAAGAAGATCGGCGTTGATGAG 382

Search completed: November 12, 2002, 16:59:27
Job time : 51.1607 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:56:36 ; Search time 27.8936 Seconds

(without alignments) 2286.857 Million cell updates/sec

Title: US-09-697-123B-22

Sequence score: 208

Scoring table: IDENTITY NUC

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/1na/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/1na/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/1na/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/1na/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/1na/PCITUS.COMB.seq.*
6: /cgn2_6/ptodata/1/1na/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	206.4	99.2	970	1	US-08-250-030-1
2	206.4	99.2	970	5	PCT-US95-06790-1
3	147.2	70.8	3447	2	US-08-313-185-57
4	147.2	70.8	3447	3	US-08-313-185-57
5	38.8	18.7	1412	1	US-08-082-614A-57
6	37	17.8	6085	4	US-09-029-603-4
7	36	17.3	1894	4	US-09-029-603-4
8	34.6	16.6	1621	1	US-08-722-001-13
9	34.6	16.6	1776	1	US-08-722-001-29
10	34.6	16.6	2140	1	US-08-314-698-1
11	34.6	16.6	2140	1	US-08-228-932-1
12	34.6	16.6	2140	1	US-08-468-939-1
13	34.6	16.6	2140	2	US-08-406-855A-1
14	34.6	16.6	2140	2	US-08-722-190-1
15	34.6	16.6	2140	3	US-08-244-354-1
16	34.6	16.6	2140	3	US-09-206-899-1
17	34.6	16.6	2140	4	US-09-444-783-1
18	34.6	16.6	2140	4	US-09-688-415-1
19	34.6	16.6	2140	5	PCT-US95-04203-1
20	34.2	16.4	22306	4	US-09-453-702B-251
21	34	16.3	680	4	US-09-556-877-267
22	34	16.3	680	4	US-09-620-412C-267
23	32.8	15.8	30001	1	US-08-125-468-1
24	32.8	15.8	30001	2	US-08-474-933-1
25	32.6	15.7	734	4	US-09-221-017B-1070
26	32.4	15.6	38584	4	US-09-453-702B-50
27	32.4	15.6	48908	4	US-09-453-702B-137

28	32	15.4	1146	3	US-08-911-853-26	Sequence 26, App1
29	32	15.4	1146	4	US-09-479-409-26	Sequence 26, App1
30	32	15.4	1146	4	US-09-479-453-26	Sequence 26, App1
31	32	15.4	17612	3	US-08-911-853-29	Sequence 29, App1
32	32	15.4	17612	4	US-09-479-409-29	Sequence 29, App1
33	32	15.4	17612	4	US-09-479-453-29	Sequence 29, App1
34	32	15.3	12122	4	US-09-029-603-1	Sequence 1, App1
35	31.8	15.3	2639	1	US-07-952-817-8	Sequence 1, App1
36	31.8	15.3	2639	6	5210025-1	Patent No. 5210025
37	31.4	15.1	1268	3	US-08-713-569-5	Sequence 5, App1
38	31.4	15.1	1320	2	US-08-461-775-8	Sequence 8, App1
39	31.4	15.1	1320	3	US-08-021-606-8	Sequence 8, App1
40	31.4	15.1	1620	2	US-08-461-775-10	Sequence 10, App1
41	31.4	15.1	1620	3	US-09-031-606-10	Sequence 10, App1
42	31.4	15.1	2167	2	US-08-461-775-9	Sequence 9, App1
43	31.4	15.1	2167	3	US-09-031-606-9	Sequence 9, App1
44	31.4	15.1	2668	2	US-08-461-775-11	Sequence 11, App1
45	31.4	15.1	2668	3	US-09-031-606-11	Sequence 11, App1

ALIGNMENTS

```
RESULT 1
US-08-250-030-1
: Sequence 1, Application US/08250030
: Patent No. 5643723
:
: APPLICANT: Persing, David H.
: TITLE OF INVENTION: Detection of a Genetic Locus Encoding
: TITLE OF INVENTION: Resistance to Rifampin in Mycobacterial Cultures and in
: TITLE OF INVENTION: Clinical Specimens
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Schwegman, Lundberg & Woessner
: STREET: 3500 IDS Center
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/250,030
: FILING DATE: 26-MAY-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Mueiting, Ann M.
: REGISTRATION NUMBER: 33,977
: REFERENCE/DOCKET NUMBER: 150.105U51
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-339-0331
: TELEFAX: 612-339-3061
: INFORMATION FOR SEQ. ID NO. 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 970 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: US-08-250-030-1
:
: Query Match          99.2%   Score 206.4;   DB 1;   Length 970;
: Best Local Similarity 99.5%   Pred. No. 4.5e-48;
: Matches 207;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;
```

QY 1 TCAGAGGAAGGCGCTACGACCTGGCCCGGCTGCTGTATAGGTCAACAGAAGCTCG 60
DB 26 TCAGAGGAAGGCGCTACGACCTGGCCCGGCTGCTGTATAGGTCAACAGAAGCTCG 85

QY 61 GCCTCATGTGCGAGCCCATCAGTCTGTCGACGCTGACCGAAGACGTCGTGGCCA 120
|||||
Db 86 GGCCTCATGTGCGAGCCCATCAGTCTGTCGACGCTGACCGAAGACGTCGTGGCCA 145
QY 121 CCATCGAATATCTGTCGCGCTTGCGACGAGGATCAGACCATGATGCTTCCGGGGCGG 180
|||||
Db 146 CCATCGAATATCTGTCGCGCTTGCGACGAGGATCAGACCATGATGACCGCTTCCGGGGCGG 205
QY 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
|||||
Db 206 TCGAGGTGCGCGGTGGAACCGACGACAT 233

RESULT 2

PCT-US95-06790-1
; Sequence 1, Application PC/TUS9506790
; GENERAL INFORMATION:
; APPLICANT: Mayo Foundation for Medical Education and Research
; APPLICANT: and Hoffmann-La Roche Inc.
; TITLE OF INVENTION: Detection of a Genetic Locus Encoding
; TITLE OF INVENTION: Resistance to Rifampin
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06790
; FILING DATE: 26-MAY-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Raasch, Kevin W.
; REGISTRATION NUMBER: 35,651
; REFERENCE/DOCKET NUMBER: 150.105M01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
PCT-US95-06790-1

Query Match 99.2%; Score 206.4; DB 5; Length 970;
Best Local Similarity 99.5%; Pred. No. 4.5e-48;

Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGAGAGAGCGCTACGACTGCGCCCGCGTGTGCTATTAAGTCAACAAGAGAGCTCG 60
|||||
Db 26 TCAAGAGAGAGCGCTACGACTGCGCCCGCGTGTGCTATTAAGTCAACAAGAGAGCTCG 85
QY 61 GCGTCATGTGCGGAGCCCATCAGTGTGTCAGCCTTACCGAAGAGACGTCGTGGCCA 120
|||||
Db 86 GCGTCATGTGCGGAGCCCATCAGTGTGTCAGCCTTACCGAAGAGACGTCGTGGCCA 145
QY 121 CCATCGAATATCTGTCGCGCTTGCGACGAGGATCAGACCATGATGCTTCCGGGGCGG 180
|||||
Db 146 CCATCGAATATCTGTCGCGCTTGCGACGAGGATCAGACCATGATGCTTCCGGGGCGG 205
QY 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
|||||
Db 206 TCGAGGTGCGCGGTGGAACCGACGACAT 233

RESULT 3

US-08-313-185-57
; Sequence 57, Application US/08313185
; Patent No. 5851763
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying
; APPLICANT: Honore, Nadine
; APPLICANT: Telenti, Amalio
; APPLICANT: Bodmer, Thomas
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
; TITLE OF INVENTION: In Mycobacterium Tuberculosis
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,185
; FILING DATE: 12-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356.0068-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 400-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-313-185-57

Query Match 70.8%; Score 147.2; DB 2; Length 3447;
Best Local Similarity 81.7%; Pred. No. 9.2e-32;

Matches 170; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1 TCAAGAGAGAGCGCTACGACTGCGCCCGCGTGTGCTATTAAGTCAACAAGAGAGCTCG 60
|||||
Db 809 TCAAGAGAGAGCGCTACGACTGCGCCCGCGTGTGCTATTAAGTCAACAAGAGAGCTCG 868
QY 61 GCGTCATGTGCGGAGCCCATCAGTGTGTCAGCCTTACCGAAGAGACGTCGTGGCCA 120
|||||
Db 869 GCGTCATGTGCGGAGCCCATCAGTGTGTCAGCCTTACCGAAGAGACGTCGTGGCCA 928
QY 121 CCATCGAATATCTGTCGCGCTTGCGACGAGGATCAGACCATGATGCTTCCGGGGCGG 180
|||||
Db 929 CCATCGAATATCTGTCGCGCTTGCGACGAGGATCAGACCATGATGCTTCCGGGGCGG 988
QY 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
|||||
Db 989 TAGAGTGCACAGTGAACCTACGATAT 1016

RESULT 4

US-09-082-614A-57

Sequence 57, Application US/09082614A
Patent No. 6124098
GENERAL INFORMATION:
APPLICANT: Heyn, Beate
APPLICANT: Cole, Stewart
APPLICANT: Young, Douglas
APPLICANT: Zhang, Ying
APPLICANT: Honore, Nadine
APPLICANT: Telenti, Amalio
APPLICANT: Bodmer, Thomas
TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
TITLE OF INVENTION: In Mycobacterium Tuberculosis
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: Dinner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,614A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/313,185
FILING DATE: 12-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 02556, 0068-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4400
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 3447 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-082-614A-57

Query Match 70.8%; Score 147.2; DB 3; Length 3447;
Best Local Similarity 81.7%; Pred. No. 9,2e-32;
Matches 170; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1 TCAGAGAAAGCGCTACGACCTGGCCGCTGCTGCTATTAAGTCAACAAGAAAGCTCG 60
DB 809 TCAGAGAAAGCGCTACGACCTGGCCGCTGCTGCTTACAAAGTCAACAAGAAAGCTCG 868
QY 61 GCGTGCATGTCGGCGAGCCATCAGTCGTGACGCTGACCGAAGAAAGAACTGTCGCA 120
DB 869 GGTTCGACCGCGGTAGTGATGATCAGTCGTGACCGCTACCGAAGAGAGATGTCGTCGA 928
QY 121 CCATCGAATATCTGCTGCGCTTGCACAGAGGTCACAGCAGCATGCTTCGCGCGCGC 180
DB 929 CCATAGAGTACCTGCTGCTGATGAGGGTCAGTCAACAATGACTGTCCAGAGTGGGG 988
QY 181 TCGAGGTGCGGTGGAACCGCAGCAT 208
DB 989 TAGAAGTCCAGTGAACCTGACATAT 1016

RESULT 5
US-08-097-831-1
Sequence 1, Application US/08097831

Patent No. 5510473
GENERAL INFORMATION:
APPLICANT: Camerini-Otero, Rafael D.
APPLICANT: Angov, Evangelina
TITLE OF INVENTION: Cloning and Expression of Taq recA
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBE, MARTENS, OLSON AND BEAR
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,831
FILING DATE: 19930726
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E.
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: NIH066,001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1412 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
IMMEDIATE SOURCE:
CLONE: Taq recA
FEATURE:
NAME/KEY: CDS
LOCATION: 62..1082
US-08-097-831-1

Query Match 18.7%; Score 38.8; DB 1; Length 1412;
Best Local Similarity 54.1%; Pred. No. 0.039;
Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 50 CAGAAGCTCGGCGTGCATGTCGGGAGCCCATCATCAGTCGTGACGCTGACCGAAGAGA 109
DB 370 CAGAAGCTCGGCGTGCATGTCGGGAGCTGCAGAGCTTCTGCTCCAGCCGACACCGGGAGCA 429
QY 110 CGTGTGCGCCACCATGATATATGTCGCTGTCAGCAGAGGTCAGACCAAGATGATCGT 169
DB 430 GCGTTTGAGATCTGTCGAGACTTCTGCGCCCTCGGGGCGGTGACGTGATCGTGTGA 489
QY 170 TCGGGGCGGCTGCGAGTCCCGTGG 195
DB 490 TTCGTGTGCGCCCTTGTGTCACCAAG 515

RESULT 6
US-09-029-603-4
Sequence 4, Application US/09029603
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Engel, Natalie
APPLICANT: Bietenader, Jurg
APPLICANT: Toupet, Christine
APPLICANT: Pospiech, Andreas
TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters


```

: GENERAL INFORMATION:
: APPLICANT: Thompson, Wayne J.
: APPLICANT: Huff, Joel R.
: APPLICANT: Nerenberg, Jennie B.
: APPLICANT: Lee, Hee-Yoon
: APPLICANT: Bell, Ian M.
: TITLE OF INVENTION: ALPHA1C ADRENERGIC RECEPTOR ANTAGONISTS
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Merck & Co., Inc.
: STREET: 126 Lincoln Avenue
: CITY: Rahway
: STATE: New Jersey
: COUNTRY: United States of America
: ZIP: 07065
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/722,001
: FILING DATE:
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/229,276
: FILING DATE: 14-APR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Appollina, Mary A.
: REGISTRATION NUMBER: 34,087
: REFERENCE/DOCKET NUMBER: 19169Y
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908)594-3462
: TELEFAX: (908)594-4720
: TELEX: 138825
: INFORMATION FOR SEQ. ID NO.: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1621 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: both
: MOLECULE TYPE: cDNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: US-08-722-001-13

Query Match          16.6%; Score 34.6; DB 1; Length 1621;
Best Local Similarity 50.3%; Pred. No. 0.57;
Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 40 ATAAAGTCAACAAGAAAGCTCGGGCTGCATGTCGGCGAGCCCATCAGTCGTGACGCTGA 99
DB 312 AGACCGTCAACCACTATTTCATGCTGAACCTGGCCGTGGCGACCTGCTGAGCGCCA 371
QY 100 CCGAAGAAGACGTGCGTGGCCACCATGCAATATCTGTCGCGTGGACGAGGCTGAGACCA 159
DB 372 CCGTACTGCCCTTCTGCGCCACCATGAGAGTTCTGGGCTTCTGGGCTTTGGCGCGCT 431
QY 160 CGATGATCGTTCGGGGCGGCGTGAAGTCCCGGTGAGAAACGACGACAT 208
DB 432 TCTGCGACGTATGGCGCCGCGTGACGTGCTGCTGCACGCGCTTCAT 480

RESULT 9
US-08-722-001-29
: Sequence 29, Application US/08722001
: GENERAL INFORMATION:
: APPLICANT: Thompson, Wayne J.
: APPLICANT: Huff, Joel R.
: APPLICANT: Nerenberg, Jennie B.
: APPLICANT: Lee, Hee-Yoon
: APPLICANT: Bell, Ian M.
```

```

: TITLE OF INVENTION: ALPHA1C ADRENERGIC RECEPTOR ANTAGONISTS
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Merck & Co., Inc.
: STREET: 126 Lincoln Avenue
: CITY: Rahway
: STATE: New Jersey
: COUNTRY: United States of America
: ZIP: 07065
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/722,001
: FILING DATE:
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/229,276
: FILING DATE: 14-APR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Appollina, Mary A.
: REGISTRATION NUMBER: 34,087
: REFERENCE/DOCKET NUMBER: 19169Y
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908)594-3462
: TELEFAX: (908)594-4720
: TELEX: 138825
: INFORMATION FOR SEQ. ID NO.: 29:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1776 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: both
: MOLECULE TYPE: cDNA
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: US-08-722-001-29

Query Match          16.6%; Score 34.6; DB 1; Length 1776;
Best Local Similarity 50.3%; Pred. No. 0.58;
Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 40 ATAAAGTCAACAAGAAAGCTCGGGCTGCATGTCGGCGAGCCCATCAGTCGTGACGCTGA 99
DB 440 AGACCGTCAACCACTATTTCATGCTGAACCTGGCCGTGGCGACCTGCTGAGCGCCA 499
QY 100 CCGAAGAAGACGTGCGTGGCCACCATGCAATATCTGTCGCGTGGACGAGGCTGAGACCA 159
DB 500 CCGTACTGCCCTTCTGCGCCACCATGAGAGTTCTGGGCTTCTGGGCTTTGGCGCGCT 559
QY 160 CGATGATCGTTCGGGGCGGCGTGAAGTCCCGGTGAGAAACGACGACAT 208
DB 560 TCTGCGACGTATGGCGCCGCGTGACGTGCTGCTGCACGCGCTTCAT 608

RESULT 10
US-08-334-698-1
: Sequence 1, Application US/08334698
: GENERAL INFORMATION:
: APPLICANT: Jonathan A. Bard et al.
: TITLE OF INVENTION: DNA Encoding Human Alpha 1 Adrenergic
: TITLE OF INVENTION: Receptors and Uses Thereof
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: COOPER & DUNHAM
: STREET: 30 Rockefeller Plaza
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10112
```

```

:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/334,698
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/952,798
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 376901
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 977-9550
: TELEFAX: (212) 664-0525
: TELEX: (212) 422523 COOP UI
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2140 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEICAL: N
: ANTI-SENSE: N
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 178..1893
: OTHER INFORMATION:
:
: US-08-334-698-1
:
: Query Match 16.6%; Score 34.6; DB 1; Length 2140;
: Best Local Similarity 50.3%; Pred. No. 0.6;
: Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
:
: QY 40 ATAAGTCAACAGAACCTGGGCTGCATGTCGCGAGCCCATCAGTCGTGACGCTGA 99
: Db 560 AGACCGTCACCACTATTTCATGTCGACCTGGCGCCACCTGCTGACGCGCA 619
: QY 100 CCGAAGAAAGCGTCGCGCCACCATGCAATATCTGTCGCTTGCACGAGGTCAGACCA 159
: Db 620 CCGTACTGCCCTTCTCGGCCACCATGAGGTTCTGGGCTTGGGCGCTTGGCGCGCT 679
: QY 160 CGATGATCTTCCGCGCGCGCTGAGTGCGCGTGAACCGACGACAT 208
: Db 680 TCTCGACGTATGGCGCGCGCTGAGCTGCTGTCTGCACGCGCTTCAT 728
:
: RESULT 11
: US-08-228-932-1
: Sequence 1, Application US/08228932
: Patent No. 5578611
: GENERAL INFORMATION:
: APPLICANT: Charles Gluchowski, Carlos C. Ferray, George Chiu,
: APPLICANT: Theresa A. Branche, John M. Weizel and Paul R. Hartig
: TITLE OF INVENTION: USE OF ALPHA-1C SPECIFIC COMPOUNDS TO TREAT BENIGN
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: COOPER & DUNHAM
: STREET: 30 Rockefeller Plaza
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10112
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/468,939
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 41878-B/JPM/TEP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 977-9550
: TELEFAX: (212) 664-0525
: TELEX: (212) 422523 COOP UI
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2140 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEICAL: N
: ANTI-SENSE: N
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 178..1893
: OTHER INFORMATION:
:
: US-08-228-932-1
:
: Query Match 16.6%; Score 34.6; DB 1; Length 2140;
: Best Local Similarity 50.3%; Pred. No. 0.6;
: Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
:
: QY 40 ATAAGTCAACAGAACCTGGGCTGCATGTCGCGAGCCCATCAGTCGTGACGCTGA 99
: Db 560 AGACCGTCACCACTATTTCATGTCGACCTGGCGCCACCTGCTGACGCGCA 619
: QY 100 CCGAAGAAAGCGTCGCGCCACCATGCAATATCTGTCGCTTGCACGAGGTCAGACCA 159
: Db 620 CCGTACTGCCCTTCTCGGCCACCATGAGGTTCTGGGCTTGGGCGCTTGGCGCGCT 679
: QY 160 CGATGATCTTCCGCGCGCGCTGAGTGCGCGTGAACCGACGACAT 208
: Db 680 TCTCGACGTATGGCGCGCGCTGAGCTGCTGTCTGCACGCGCTTCAT 728
:
: RESULT 12
: US-08-468-939-1
: Sequence 1, Application US/08468939
: Patent No. 5714381
: GENERAL INFORMATION:
: APPLICANT: Jonathan A. Bard et al.
: APPLICANT: DNA Encoding Human Alpha 1 Adrennergic
: TITLE OF INVENTION: Receptors and Uses Thereof
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: COOPER & DUNHAM LLP
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/468,939
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
```

```

:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/468,939
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
```

```

:
: SOFTWARE: Patentin Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/228,932
: FILING DATE: 13-Apr-1994
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 41878-B/JPM/TEP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 977-9550
: TELEFAX: (212) 664-0525
: TELEX: (212) 422523 COOP UI
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2140 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEICAL: N
: ANTI-SENSE: N
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 178..1893
: OTHER INFORMATION:
:
: US-08-228-932-1
:
: Query Match 16.6%; Score 34.6; DB 1; Length 2140;
: Best Local Similarity 50.3%; Pred. No. 0.6;
: Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
:
: QY 40 ATAAGTCAACAGAACCTGGGCTGCATGTCGCGAGCCCATCAGTCGTGACGCTGA 99
: Db 560 AGACCGTCACCACTATTTCATGTCGACCTGGCGCCACCTGCTGACGCGCA 619
: QY 100 CCGAAGAAAGCGTCGCGCCACCATGCAATATCTGTCGCTTGCACGAGGTCAGACCA 159
: Db 620 CCGTACTGCCCTTCTCGGCCACCATGAGGTTCTGGGCTTGGGCGCTTGGCGCGCT 679
: QY 160 CGATGATCTTCCGCGCGCGCTGAGTGCGCGTGAACCGACGACAT 208
: Db 680 TCTCGACGTATGGCGCGCGCTGAGCTGCTGTCTGCACGCGCTTCAT 728
:
: RESULT 12
: US-08-468-939-1
: Sequence 1, Application US/08468939
: Patent No. 5714381
: GENERAL INFORMATION:
: APPLICANT: Jonathan A. Bard et al.
: APPLICANT: DNA Encoding Human Alpha 1 Adrennergic
: TITLE OF INVENTION: Receptors and Uses Thereof
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: COOPER & DUNHAM LLP
: STREET: 1185 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/468,939
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John P.
: REGISTRATION NUMBER: 28,678
```


OTHER INFORMATION:
US-08-722-190-1

Query Match 16.6%; Score 34.6; DB 2; Length 2140;
Best Local Similarity 50.3%; Pred. No. 0.6;
Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 40 ATAGGTCACAGAGCTGGGCTCATGTCGGCAGCCCATCATGCTGACGCTGA 99
DB 560 AGACCGTCACCACTATTTCATGCTGAACCTGGCGCTGGCCACTGCTGTGAGCGCCA 619
QY 100 CCGAAGAGAGCTGCTGGCCACCATGAAATATCTGCTGACGAGGCTGAGACCA 159
DB 620 CCGTACTGCCCTTCTGGCCACCATGAGAGTCTGGGCTTGTGGCCCTTGGCGCGCT 679
QY 160 CGATGATCGTTCCGGCGGCTGAGGTCCCGTGAAACGACGACAT 208
DB 680 TCTGCGACGTATGGCGCGCTGAGAGTCTGCTGCACGCGCTCCAT 728

RESULT 15

US-08-244-354-1
Sequence 1, Application US/08244354
Patent No. 6015819

GENERAL INFORMATION:

APPLICANT: Charles Gluchowski, et al.
TITLE OF INVENTION: USE OF ALPHA-1C SPECIFIC COMPOUNDS TO
TREAT BENIGN PROSTATIC HYPERPLASIA
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:

ADDRESSEE: COOPER & DUNHAM LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/244,354

FILING DATE: April 1, 1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2140 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: N

FEATURE:

NAME/KEY: CDS

LOCATION: 178..1893

OTHER INFORMATION:

US-08-244-354-1

Query Match 16.6%; Score 34.6; DB 3; Length 2140;

Best Local Similarity 50.3%; Pred. No. 0.6;

Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 40 ATAGGTCACAGAGCTGGGCTCATGTCGGCAGCCCATCATGCTGACGCTGA 99

DB 560 AGACCGTCACCACTATTTCATGCTGAACCTGGCGCTGGCCACTGCTGTGAGCGCCA 619
QY 100 CCGAAGAGAGCTGCTGGCCACCATGAAATATCTGCTGACGAGGCTGAGACCA 159
DB 620 CCGTACTGCCCTTCTGGCCACCATGAGAGTCTGGGCTTGTGGCCCTTGGCGCGCT 679
QY 160 CGATGATCGTTCCGGCGGCTGAGGTCCCGTGAAACGACGACAT 208
DB 680 TCTGCGACGTATGGCGCGCTGAGAGTCTGCTGCACGCGCTCCAT 728

Search completed: November 12, 2002, 21:03:40
Job time: 39.8936 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 ; Search time 136.222 Seconds

(without alignments) 3438.621 Million cell updates/sec

Title: US-09-697-123b-22

Perfect score: 208

Sequence: 1 tcaagagaagcgtacgac.....ccgttgaaaccagcagacat 208

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	208	100.0	208	22	AA505222
2	206.4	99.2	208	22	AA505205
3	206.4	99.2	208	22	AA505206
4	206.4	99.2	208	22	AA505216
5	206.4	99.2	970	17	AA709676
6	206.4	99.2	3519	22	AAH51976
7	206.4	99.2	3534	22	AAH02079
8	206.4	99.2	3853	21	AAH74651
9	206.4	99.2	3853	21	AAH89994

10	171.2	82.3	208	22	AA505210	Mycobacterium ulce
11	169.6	81.5	208	22	AA505203	Mycobacterium gord
12	169.6	81.5	208	22	AA505211	Mycobacterium mari
13	168	80.8	208	22	AA505215	Mycobacterium aviu
14	164.8	79.2	208	22	AA505201	Mycobacterium gord
15	164.8	79.2	208	22	AA505208	Mycobacterium kans
16	164.8	79.2	208	22	AA505219	Mycobacterium fort
17	164.8	79.2	208	22	AA505223	Mycobacterium haem
18	163.2	78.5	208	22	AA505202	Mycobacterium gord
19	158.4	76.2	208	22	AA505217	Mycobacterium cela
20	157.6	75.8	207	22	AA505204	Mycobacterium gord
21	152.8	73.5	207	22	AA505212	Mycobacterium szul
22	150.4	72.3	208	22	AA505224	Mycobacterium xeno
23	147.2	70.8	3447	14	AA051532	M. lepre prob gene
24	145.6	70.0	205	22	AA505220	Mycobacterium intr
25	137.6	66.2	214	22	AA505214	Mycobacterium malin
26	136.6	65.7	223	22	AA505209	Mycobacterium scro
27	128	61.5	214	22	AA505221	Mycobacterium abs
28	124.6	59.9	211	22	AA505218	Mycobacterium flav
29	123.2	59.2	214	22	AA505207	Mycobacterium chel
30	122.2	58.8	223	22	AA505213	Mycobacterium gasc
31	88	42.3	3495	22	AAH65512	C glutamicum codin
32	88	42.3	349980	22	AAH68525	Propionibacterium
33	64.4	31.0	27426	23	AA59541	Burkholderia cepac
34	39	18.8	1041	21	AA251711	Burkholderia cepac
35	38.8	18.7	1412	17	AA251711	Burkholderia cepac
36	38.2	18.4	1041	21	AA251715	Thermus aquaticus
37	38	18.3	1119	22	AA28350	Burkholderia cepac
38	37.4	18.0	1041	21	AA251704	Streptomyces ambol
39	37.4	18.0	1041	21	AA251710	Burkholderia vietn
40	37.4	18.0	1041	21	AA251712	Burkholderia cepac
41	37.4	18.0	1041	21	AA251712	Burkholderia cepac
42	37.4	18.0	1041	21	AA251714	Burkholderia cepac
43	37.4	18.0	1044	21	AA251698	Burkholderia cepac
44	37	17.8	6085	18	AA70153	S. longisporoflavus
45	36	17.3	1894	18	AA76542	50K-cellulase gene

ALIGNMENTS

RESULT 1
AA505222
ID AA505222 standard; DNA; 208 BP.
XX
AC AA505222;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium africanum rpoB gene fragment.
XX
KM Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium africanum.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
DR WPI: 2001-300520/31.
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -

XX Claim 1; Page 47; 50pp: English.
XX
PS
CC The present sequence for *Mycobacterium africanum* rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various *Mycobacterium* species. These rpoB gene fragments can be used
CC in the diagnosis and identification of *Mycobacterium* species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of *Mycobacterium* species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including *M. tuberculosis*, *M. leprae*
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific *Mycobacterium* species.
XX
SQ Sequence 208 BP; 46 A; 61 C; 67 G; 34 T; 0 other;
Query Match 100.0%; Score 208; DB 22; Length 208;
Best Local Similarity 100.0%; Pred. No. 5, 2e-46;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TCAGAGAGAACGCTACGACCTGCGCGCTGCTGCTATTAAGTCAACAGAAAGCTCG 60
DB 1 TCAGAGAGAACGCTACGACCTGCGCGCTGCTGCTATTAAGTCAACAGAAAGCTCG 60
QY 61 GCGTCATGTCGGCGAGCCCATCACTGCTGACAGCTGACCGAAGAGAGCTGCGCCA 120
DB 61 GCGTCATGTCGGCGAGCCCATCACTGCTGACAGCTGACCGAAGAGAGCTGCGCCA 120
QY 121 CCATCGAATATCTGTCGCTTGGCAGCAGAGGTCAAGACAGATGATGTTCCGGCGCGC 180
DB 121 CCATCGAATATCTGTCGCTTGGCAGCAGAGGTCAAGACAGATGATGTTCCGGCGCGC 180
QY 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
DB 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
RESULT 2
AAS05205
ID AAS05205 standard; DNA: 208 BP.
XX
AC AAS05205;
XX
DT 07-SEP-2001 (first entry)
XX
DE *Mycobacterium tuberculosis* rpoB gene fragment.
XX
KM Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.
XX
OS *Mycobacterium tuberculosis*.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX
DR WPI: 2001-300520/31.
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism
XX
PS Disclosure; Page 41; 50pp: English.
XX
SQ The present sequence for *Mycobacterium tuberculosis* rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various *Mycobacterium* species. These rpoB gene fragments can be used
CC in the diagnosis and identification of *Mycobacterium* species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of *Mycobacterium* species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including *M. tuberculosis*, *M. leprae*
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific *Mycobacterium* species.
XX
SQ Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
Query Match 99.2%; Score 206.4; DB 22; Length 208;
Best Local Similarity 99.5%; Pred. No. 1.6e-45;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 TCAGAGAGAACGCTACGACCTGCGCGCTGCTGCTATTAAGTCAACAGAAAGCTCG 60
DB 1 TCAGAGAGAACGCTACGACCTGCGCGCTGCTGCTATTAAGTCAACAGAAAGCTCG 60
QY 61 GCGTCATGTCGGCGAGCCCATCACTGCTGACAGCTGACCGAAGAGAGCTGCGCCA 120
DB 61 GCGTCATGTCGGCGAGCCCATCACTGCTGACAGCTGACCGAAGAGAGCTGCGCCA 120
QY 121 CCATCGAATATCTGTCGCTTGGCAGCAGAGGTCAAGACAGATGATGTTCCGGCGCGC 180
DB 121 CCATCGAATATCTGTCGCTTGGCAGCAGAGGTCAAGACAGATGATGTTCCGGCGCGC 180
QY 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
DB 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
RESULT 3
AAS05206
ID AAS05206 standard; DNA: 208 BP.
XX
AC AAS05206;
XX
DT 07-SEP-2001 (first entry)
XX
DE *Mycobacterium terrae* rpoB gene fragment.
XX
KM Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.
XX
OS *Mycobacterium terrae*.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.

```
XX 27-OCT-1999; 99KR-0046795.
XX (ERUM-) ERUME BIOTECH CO LTD.
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ:
XX WP1; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism
XX
XX Claim 1; Page 42; 50pp; English.
XX
XX The present sequence for Mycobacterium terrae rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX selection of appropriate epidemiological and pathogenesis information for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other:
XX
XX Query Match 99.2%; Score 206.4; DB 22; Length 208;
XX Best Local Similarity 99.5%; Pred. No. 1.6e-45;
XX Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
OY 1 TCAAGGAGAACGGCTACGACCTGCGCGCGTGGTCTATTAAGCTCAACAAGACCTCG 60
DB 1 TCAAGGAGAACGGCTACGACCTGCGCGCGTGGTCTATTAAGCTCAACAAGACCTCG 60
OY 61 GGCTGCATGTCGGCGAGCCCATCAGCTGCTGACGCTGACCGAAGAAGAGCTGTGGCCA 120
DB 61 GGCTGCATGTCGGCGAGCCCATCAGCTGCTGACGCTGACCGAAGAAGAGCTGTGGCCA 120
OY 121 CCATCGAATATCTGTCGCTTGCACGAGGTCAGACGACGATGATCGTTCCGGGCGCG 180
DB 121 CCATCGAATATCTGTCGCTTGCACGAGGTCAGACGACGATGATCGTTCCGGGCGCG 180
OY 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
DB 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
XX
XX RESULT 4
XX AAS05216
XX ID AAS05216 standard; DNA; 208 BP.
XX AC AAS05216;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium bovis rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
XX Mycobacterium bovis.
XX
```

```
XX WO20011061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ:
XX WP1; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism
XX
XX Claim 1; Page 45; 50pp; English.
XX
XX The present sequence for Mycobacterium bovis rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX selection of appropriate epidemiological and pathogenesis information for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other:
XX
XX Query Match 99.2%; Score 206.4; DB 22; Length 208;
XX Best Local Similarity 99.5%; Pred. No. 1.6e-45;
XX Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
OY 1 TCAAGGAGAACGGCTACGACCTGCGCGCGTGGTCTATTAAGCTCAACAAGACCTCG 60
DB 1 TCAAGGAGAACGGCTACGACCTGCGCGCGTGGTCTATTAAGCTCAACAAGACCTCG 60
OY 61 GGCTGCATGTCGGCGAGCCCATCAGCTGCTGACGCTGACCGAAGAAGAGCTGTGGCCA 120
DB 61 GGCTGCATGTCGGCGAGCCCATCAGCTGCTGACGCTGACCGAAGAAGAGCTGTGGCCA 120
OY 121 CCATCGAATATCTGTCGCTTGCACGAGGTCAGACGACGATGATCGTTCCGGGCGCG 180
DB 121 CCATCGAATATCTGTCGCTTGCACGAGGTCAGACGACGATGATCGTTCCGGGCGCG 180
OY 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
DB 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
XX
XX RESULT 5
XX AAT09676
XX ID AAT09676 standard; DNA; 970 BP.
XX AC AAT09676;
XX
XX 15-OCT-1996 (first entry)
XX
```

```

DE   Mycobacterium tuberculosis rpoB gene DNA sequence.
XX
KW   Tuberculosis: disease diagnosis: Oligonucleotide: DNA primer: PCR;
KW   polymerase chain reaction: DNA amplification: rpoB locus: TB; ss.
XX
OS   Mycobacterium tuberculosis.
XX
FH   Key
FH   10..27
FH   primer_bind
FH   /tag= a
FH   /note= "primer FENLEP"
FH   226..243
FH   primer_bind
FH   /tag= b
FH   /note= "primer DDIDL"
FH   226..240
FH   primer_bind
FH   /tag= c
FH   /note= "primer DDIDH"
FH   338..364
FH   primer_bind
FH   /tag= d
FH   /note= "primer rpo95"
FH   348..373
FH   primer_bind
FH   /tag= e
FH   /note= "primer rpo105"
FH   354..373
FH   primer_bind
FH   /tag= f
FH   /note= "primer KY290"
FH   372..373
FH   misc_feature
FH   /tag= g
FH   /note= "M. tuberculosis signature nucleotide"
FH   433..434
FH   misc_feature
FH   /tag= h
FH   /note= "M. tuberculosis signature nucleotide"
FH   438
FH   misc_feature
FH   /tag= i
FH   /note= "M. tuberculosis signature nucleotide"
FH   468..469
FH   misc_feature
FH   /tag= j
FH   /note= "M. tuberculosis signature nucleotide"
FH   486
FH   misc_feature
FH   /tag= k
FH   /note= "M. tuberculosis signature nucleotide"
FH   501
FH   misc_feature
FH   /tag= l
FH   /note= "M. tuberculosis signature nucleotide"
FH   516
FH   misc_feature
FH   /tag= m
FH   /note= "M. tuberculosis signature nucleotide"
FH   516..535
FH   primer_bind
FH   /tag= n
FH   /note= "primer rpo273"
FH   525
FH   misc_feature
FH   /tag= o
FH   /note= "M. tuberculosis signature nucleotide"
FH   525..541
FH   primer_bind
FH   /tag= p
FH   /note= "primer KY292"
FH   536..562
FH   primer_bind
FH   /tag= q
FH   /note= "primer rpo293"
FH   640..666
FH   primer_bind
FH   /tag= r
FH   /note= "primer rpo397"
FH   952..966
FH   primer_bind
FH   /tag= s
FH   /note= "primer NMQRO-1"
FH   952..966
FH   primer_bind
FH   /tag= t
FH   /note= "primer NMQRO-2"
XX
XX   WO9533074-A1.
XX
XX   07-DEC-1995.
XX

```

```

PF   26-MAY-1995; 95WO-0506790.
XX
XX   26-MAY-1994; 94US-0250030.
XX
XX   (HOFF ) HOFFMANN LA ROCHE INC.
XX   (MAYO-) MAYO FOUNDATION.
XX
XX   Felmlae TA, Hunt JM, Persing DH, Roberts GD, Whelen AC;
XX   Young KKY;
XX
XX   WPI; 1996-030581/03.
XX
XX   Detection of Mycobacterium tuberculosis - by amplifying sample DNA
XX   with a primer set that targets portions of the gene encoding rpoB.
XX
XX   Disclosure; Fig.3; 54pp; English.
XX

```

This oligonucleotide DNA primer is specific for Mycobacterium tuberculosis, and may be used to amplify a sample DNA by targeting a portion of the gene encoding rpoB. The 1st several bases comprise a nonhybridizing tail consisting of filler bases followed by a restriction site incorporated to facilitate cloning using the CC amplicon at a later date, if desired. The remaining bases hybridize to bacterial rpoB DNA. The method provides for the detection of M. tuberculosis and the concurrent determination of its drug susceptibility, particularly to rifampicin. The method can provide often greater than 95% sensitivity and 100% specificity. The biological sample is a fluid or tissue sample from a human.

Sequence 970 BP; 182 A; 302 C; 330 G; 156 T; 0 other;

Query Match 99.2%; Score 206.4; DB 17; Length 970;
 Best Local Similarity 99.5%; Pred. No. 2e-45;
 Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY   1 TCAAGAGAAGCGCTGACGACCTGGCCCGGTGGCTATTAAGTCAACAAGAGCTCG 60
DB   |||||||
QY   26 TCAAGAGAAGCGCTGACGACCTGGCCCGGTGGCTATTAAGTCAACAAGAGCTCG 85
DB   |||||||
QY   61 GGCTGATGTCGGCGAGCCCATCAGTCGTCGACGCTGACCGAAGAGCTGTGGCCA 120
DB   |||||||
QY   86 GGCTGATGTCGGCGAGCCCATCAGTCGTCGACGCTGACCGAAGAGCTGTGGCCA 145
DB   |||||||
QY   121 CCATCGATATGTCGGCTGGCCCTGCACAGAGGTCAACACCATGATGTCGGCGCGC 180
DB   |||||||
QY   146 CCATCGATATGTCGGCTGGCCCTGCACAGAGGTCAACACCATGATGTCGGCGCGC 205
DB   |||||||
QY   181 TCGAGGTGCGGTGGAACCGACGACAT 208
DB   |||||||
QY   206 TCGAGGTGCGGTGGAACCGACGACAT 233
DB   |||||||

```

```

RESULT 6
AAH51976
ID   AAH51976 standard; DNA; 3519 BP.
XX
XX   AAH51976;
XX
XX   04-SEP-2001 (first entry)
XX
XX   Mycobacterium tuberculosis potential drug target gene SEQ ID 30.
XX
XX   Drug target; growth; organism viability; characterisation; ds.
XX
XX   Mycobacterium tuberculosis.
XX
XX   WO200135317-A1.
XX
XX   17-MAY-2001.
XX
XX   13-NOV-2000; 2000WO-US31152.
XX
XX   12-NOV-1999; 99US-0165086.
XX

```



```
ID    AAA74651 standard; DNA; 3853 BP.
XX
XX
AC    AAA74651;
XX
XX
DT    06-DEC-2000 (first entry)
XX
XX
DE    Mycobacterium tuberculosis rpoB gene.
XX
XX
DE    Mycobacterium tuberculosis; rpoB; RNA polymerase beta subunit;
XX
XX
KM    rifampin resistance; mutation detection; ds.
XX
XX
OS    Mycobacterium tuberculosis.
XX
XX
PN    WO200043546-A2.
XX
XX
PD    27-JUL-2000.
XX
XX
PF    20-DEC-1999; 99MO-US30377.
XX
XX
PR    19-JAN-1999; 99US-0233996.
XX
XX
PA    (DADE-) DADE BEHRING INC.
XX
XX
PI    Liu YP, Kurn N;
XX
XX
DR    WPI: 2000-524243/47.
XX
XX
PT    Method for detecting drug resistance in a strain of an organism,
XX
XX
PT    particularly for detecting rifampin resistance in Mycobacterium
XX
XX
PS    tuberculosis -
XX
XX
PS    Example 1; Fig 4; 86bp; English.
XX
XX
CC    The present sequence is the rpoB gene from Mycobacterium tuberculosis.
CC
CC    Rifampin resistance is largely associated with point mutations
CC    localised in a small core region of 81 base pairs in the rpoB gene, which
CC    encodes the RNA polymerase beta subunit. To detect a mutation, a complex
CC    is formed comprising a first sequence representing the predetermined
CC    region of the gene of the organism and a second sequence representing the
CC    corresponding region of the gene of the wild type organism in double
CC    stranded form. Each member of at least one pair of non-complementary
CC    strands within the complex has a label. The association of the labels in
CC    the complex is related to the presence of the mutation. The presence of
CC    the mutation is related to the drug resistance of the strain.
XX
XX
SQ    Sequence 3853 BP; 723 A; 1173 C; 1293 G; 664 T; 0 other;
XX
XX
Query Match          99.2%; Score 206.4; DB 21; Length 3853;
Best Local Similarity 99.5%; Pred. No. 2.3e-45;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX
QY    1 TCAAGGAGAAGCGCTACGACCTGGCCCGCGTGGTCTATTAAGTCAACAAGAAGCTCG 60
DB    1477 TCAAGGAGAAGCGCTACGACCTGGCCCGCGTGGTCTATTAAGTCAACAAGAAGCTCG 1536
XX
XX
QY    61 GGCTGATGTCGGCGAGCCATCAGCTGCTGACGCTGACCGAAGAAGAGCTGCGCCA 120
DB    1537 GGCTGATGTCGGCGAGCCATCAGCTGCTGACGCTGACCGAAGAAGAGCTGCGCCA 1596
XX
XX
QY    121 CCATCGAATATCTGTCGCTTGCACAGAGGTCAAGACCAAGATGATCTTCCGGGCGCG 180
DB    1597 CCATCGAATATCTGTCGCTTGCACAGAGGTCAAGACCAAGATGATCTTCCGGGCGCG 1656
XX
XX
QY    181 TCGAGGTGCCGGTGAACCGACGACAT 208
DB    1657 TCGAGGTGCCGGTGAACCGACGACAT 1684
XX
XX
RESULT 9
ID    AAA89994 standard; DNA; 3853 BP.
XX
XX
AC    AAA89994;
```

```
XX
XX
DT    18-DEC-2000 (first entry)
XX
XX
DE    M. tuberculosis rpoB gene encoding RNA polymerase beta subunit.
XX
XX
DE    Drug resistance; rifampin; rifampicin; pyrazinamide; rpoB;
XX
XX
KM    RNA polymerase beta subunit; ds.
XX
XX
OS    Mycobacterium tuberculosis.
XX
XX
PN    WO200043545-A2.
XX
XX
PD    27-JUL-2000.
XX
XX
PF    14-DEC-1999; 99MO-US29517.
XX
XX
PR    19-JAN-1999; 99US-0233996.
XX
XX
PR    22-APR-1999; 99US-0296894.
XX
XX
PA    (DADE-) DADE BEHRING INC.
XX
XX
PI    Liu YP, Kurn N;
XX
XX
DR    WPI: 2000-499235/44.
XX
XX
PT    Detecting resistance of drugs such as rifampicin in strains of
XX
XX
PT    Mycobacterium, comprising detecting mutations in a gene and relating
XX
XX
PS    them to drug resistance -
XX
XX
PS    Example 1; Fig 4; 91bp; English.
XX
XX
CC    This invention relates to a method for detecting drug resistance in a
CC    strain of an organism. The method comprises detecting the presence of at
CC    least 1 mutation in a first sequence and relating the presence of the
CC    mutation to drug resistance. Included in the invention are a kit for
CC    carrying out the method and a method for detecting the presence of a
CC    difference between two related nucleic acid sequences in an organism. The
CC    methods are useful for detecting resistance to drugs such as rifampin and
CC    pyrazinamide in Mycobacterium. The present sequence represents the
CC    Mycobacterium tuberculosis rpoB gene (which encodes the RNA polymerase
CC    beta subunit). The sequence is used in an example of the method of the
CC    invention for the detection of rifampin resistance in M. tuberculosis.
XX
XX
SQ    Sequence 3853 BP; 723 A; 1173 C; 1293 G; 664 T; 0 other;
XX
XX
Query Match          99.2%; Score 206.4; DB 21; Length 3853;
Best Local Similarity 99.5%; Pred. No. 2.3e-45;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX
QY    1 TCAAGGAGAAGCGCTACGACCTGGCCCGCGTGGTCTATTAAGTCAACAAGAAGCTCG 60
DB    1477 TCAAGGAGAAGCGCTACGACCTGGCCCGCGTGGTCTATTAAGTCAACAAGAAGCTCG 1536
XX
XX
QY    61 GGCTGATGTCGGCGAGCCATCAGCTGCTGACGCTGACCGAAGAAGAGCTGCGCCA 120
DB    1537 GGCTGATGTCGGCGAGCCATCAGCTGCTGACGCTGACCGAAGAAGAGCTGCGCCA 1596
XX
XX
QY    121 CCATCGAATATCTGTCGCTTGCACAGAGGTCAAGACCAAGATGATCTTCCGGGCGCG 180
DB    1597 CCATCGAATATCTGTCGCTTGCACAGAGGTCAAGACCAAGATGATCTTCCGGGCGCG 1656
XX
XX
QY    181 TCGAGGTGCCGGTGAACCGACGACAT 208
DB    1657 TCGAGGTGCCGGTGAACCGACGACAT 1684
XX
XX
RESULT 10
ID    AAS05210 standard; DNA; 208 BP.
XX
XX
AC    AAS05210;
XX
XX
DT    07-SEP-2001 (first entry)
```


XX DE Mycobacterium ulcerans rpoB gene fragment.
XX KM Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX KM PCR-restriction fragment length polymorphism analysis; ds.
XX OS Mycobacterium ulcerans.
XX PN WO200131061-A1.
XX PD 03-MAY-2001.
XX PF 27-OCT-2000; 2000MO-KR01223.
XX PR 27-OCT-1999; 99KR-0046795.
XX PA (ERUM-) ERUME BIOTECH CO LTD.
XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX DR WPI; 2001-300520/31.
XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for
XX PT diagnosis and identification of many mycobacterial species by
XX PT restriction fragment length polymorphism -
XX PS Claim 1; Page 43; 50pp; English.
XX CC The present sequence for Mycobacterium ulcerans rpoB gene
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX CC various Mycobacterial species. These rpoB gene fragments can be used
XX CC in the diagnosis and identification of Mycobacterium species using a
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)
XX CC method. The method comprises obtaining a restriction fragment length
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX CC amplifying and digesting the DNA fragment from the microorganism to
XX CC be identified and comparing the RFLP patterns from the known rpoB gene
XX CC fragments with the unidentified fragment. The rpoB gene fragments
XX CC are useful to identify a wide range of Mycobacterium species, e.g. for
XX CC diagnosis or to obtain epidemiological and pathogenesis information for
XX CC selection of appropriate therapies, including M. tuberculosis, M. lepre
XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR
XX CC required), and can differentiate between many species in a single
XX CC experiment, including those difficult to distinguish by usual biochemical
XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX CC detecting specific Mycobacterial species.
SQ Sequence 208 BP; 45 A; 68 C; 67 G; 28 T; 0 other;
Query Match 82.3%; Score 171.2; DB 22; Length 208;
Best Local Similarity 88.9%; Pred. No. 3.4e-36;
Matches 185; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
0Y 1 TCAAGGAGAGCGGTACGACTGGCCGCTGCTGCTTAAGTCAACAGACCTCG 60
DB 1 TCAAGGAGAGCGGTACGACTGGCTGCTGCTGCTGCTCAAGGTCAACAGACCTCG 60
0Y 61 GGCTGATGTGCGGAGCCCATCATCGTCTGACGCTGACCGAAGAAAGCTGGCCA 120
DB 61 GGCTGAGAACGGCGGCGAGCCATCATCGTCTGACGCTGACCGAAGAAAGCTGGCCA 120
0Y 121 CCATCGAATATCTGCTCCGCTTGACGAGGGTCAAGACGATGATCTTCCGGCGGCG 180
DB 121 CCATCGAATATCTGCTCCGCTTGACGAGGGTCAAGACGATGATGACCGGGCGGCTG 180
0Y 181 TCGAGTGCCTGGTGAAGACGAGCAT 208
DB 181 TCGAGTGCCTGGTGAAGACGAGCAT 208
RESULT 11

AAS05203
ID AAS05203 standard; DNA; 208 BP.
XX AAS05203;
AC AAS05203;
DT 07-SEP-2001 (first entry)
XX DE Mycobacterium gordonae type III rpoB gene fragment.
XX KM Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX KM PCR-restriction fragment length polymorphism analysis; ds.
XX OS Mycobacterium gordonae type III.
XX PN WO200131061-A1.
XX PD 03-MAY-2001.
XX PF 27-OCT-2000; 2000MO-KR01223.
XX PR 27-OCT-1999; 99KR-0046795.
XX PA (ERUM-) ERUME BIOTECH CO LTD.
XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX DR WPI; 2001-300520/31.
XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for
XX PT diagnosis and identification of many mycobacterial species by
XX PT restriction fragment length polymorphism -
XX PS Claim 1; Page 41; 50pp; English.
XX CC The present sequence for Mycobacterium gordonae type III rpoB gene
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX CC various Mycobacterial species. These rpoB gene fragments can be used
XX CC in the diagnosis and identification of Mycobacterium species using a
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)
XX CC method. The method comprises obtaining a restriction fragment length
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX CC amplifying and digesting the DNA fragment from the microorganism to
XX CC be identified and comparing the RFLP patterns from the known rpoB gene
XX CC fragments with the unidentified fragment. The rpoB gene fragments
XX CC are useful to identify a wide range of Mycobacterium species, e.g. for
XX CC diagnosis or to obtain epidemiological and pathogenesis information for
XX CC selection of appropriate therapies, including M. tuberculosis, M. lepre
XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR
XX CC required), and can differentiate between many species in a single
XX CC experiment, including those difficult to distinguish by usual biochemical
XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX CC detecting specific Mycobacterial species.
SQ Sequence 208 BP; 46 A; 69 C; 63 G; 30 T; 0 other;
Query Match 81.5%; Score 169.6; DB 22; Length 208;
Best Local Similarity 88.5%; Pred. No. 8.9e-36;
Matches 184; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
0Y 1 TCAAGGAGAGCGGTACGACTGGCCGCTGCTGCTTAAGTCAACAGACCTCG 60
DB 1 TCAAGGAGAGCGGTACGACTGGCCGCTGCTGCTGCTCAAGGTCAACAGACCTCG 60
0Y 61 GGCTGATGTGCGGAGCCCATCATCGTCTGACGCTGACCGAAGAAAGCTGGCCA 120
DB 61 GGCTGACACGTGCGGAGATCCATCATCGTCTGACGCTGACCGAAGAAAGCTGGCCA 120
0Y 121 CCATCGAATATCTGCTCCGCTTGACGAGGGTCAAGACGATGATCTTCCGGCGGCG 180
DB 121 CCATCGAATATCTGCTCCGCTTGACGAGGGTCAAGACGATGATGACCGGGCGGCTG 180

OY 181 TCGAGTGCCTGGTGAACCGACGACAT 208
||||| ||||||| ||||||| |||||||
Db 181 CCGAGGTCTCCGCTGACACCGACGACAT 208

RESULT 12

AAS05211
ID AAS05211 standard; DNA: 208 BP.

AC AAS05211;

DT 07-SEP-2001 (first entry)

DE Mycobacterium marinum rpoB gene fragment.

KM Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium marinum.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI: 2001-300520/31.

DR WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism.

PS Claim 1; Page 43; 50pp; English.

XX The present sequence for Mycobacterium marinum rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 45 A; 69 C; 68 G; 26 T; 0 other;

SO Query Match 81.5%; Score 169.6; DB 22; Length 208;

Best Local Similarity 88.5%; Pred. No. 8.9e-36;

Matches 184; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

OY 1 TCAAGAGAGCGCTACGACCTGCGCGCTGCTGCTAATAGGTCAACAGAAGCTCG 60

Db 1 TCAAGAGAGAGCGCTACGACCTGCGCGCGCTGCTGCTAATAGGTCAACAGAAGCTCG 60

OY 61 GCGTCATGTCGGGCGAGCCCATCATCGTGCAGCTGACCCGAAGAAGACGTGTCGCCCA 120

Db 61 GCGTCAGACGCCGGCCACCCCATCATCAGCTGCAGACCTGACCGAGGAAGACGTGTCGCCCA 120
OY 121 CCATCGAATATCTGACCGCTGACAGAGGTGACACGACGATGATCTCCGGCGCGC 180
Db 121 CCATCGAATATCTGCTCGCTTGCACGAGGCGCCAGACCGCATGACCGCTCCGGCGCGTG 180
OY 181 TCGAGTGCCTGGTGAACCGACGACAT 208
||||| ||||||| ||||||| |||||||
Db 181 TCGAGTGCCTGGTGAACCGACGACAT 208

RESULT 13

AAS05215
ID AAS05215 standard; DNA: 208 BP.

AC AAS05215;

DT 07-SEP-2001 (first entry)

DE Mycobacterium avium rpoB gene fragment.

KM Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium avium.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI: 2001-300520/31.

DR WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism.

PS Claim 1; Page 44; 50pp; English.

XX The present sequence for Mycobacterium avium rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;

SO Query Match 80.8%; Score 168; DB 22; Length 208;

Best Local Similarity 88.0%; Pred. No. 2.4e-35;

Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

and non-tuberculous mycobacteria (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the

CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
experiment, including those difficult to distinguish by usual biochemical
tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

XX
SQ Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;

Query Match 79.2%: Score 164.8; DB 22; Length 208;
Best Local Similarity 87.0%; Pred. No. 1.7e-34;

Matches 181: Conservative 0; Mismatches 27; Indels 0; Gaps 0;

OY 1 TCACGAGAGAGCGCTACGACCTGGCCCGCGTGGCTGCTAAGGTCAACACAGAGCTGG 60
|||||
DB 1 TCACGAGAGAGCGCTACGACCTGGCCCGCGTGGCTGCTAAGGTCAACACAGAGCTGG 60
|||||
OY 61 GGCTGCATGTGGGGAGGCCATCAGCTGTCGACGCTGACCGAAGAGACGTGGCCA 120
|||||
DB 61 GGCTGCATGTGGGGAGGCCATCAGCTGTCGACGCTGACCGAAGAGACGTGGCCA 120
|||||
OY 121 CCATCGAATATCTGTGCTCGCTTGCACGAGGAGTCAACACAGATGATCGTTCCGGGGGGCGG 180
|||||
DB 121 CCATCGAATATCTGTGCTCGCTTGCACGAGGAGTCAACACAGATGATCGTTCCGGGGGGCGG 180
|||||
OY 181 TCGAGGTGCCGGTGAACCGAGACAT 208
|||||
DB 181 TCGAGGTGCCGGTGAACCGAGACAT 208
|||||

Search completed: November 12, 2002, 16:50:31
Job time : 142.222 secs

GenCore Version 5.1.3
Copyright (c) 1993 2002 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 ; Search time 636.023 Seconds

(without alignments) 9517.553 Million cell updates/sec

Title: US-09-697-123b-22

Perfect score:

208

Sequence: 1 tcaagagagagcgtacgac.....ccggtggaaccgacacat 208

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlg_hum:*
40: em_hlg_mus:*
41: em_hlg_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	206.4	99.2	610	1	MTU318818
2	206.4	99.2	610	1	MTU318819
3	206.4	99.2	616	1	MTU318813
4	206.4	99.2	618	1	MTU318815
5	206.4	99.2	618	1	MTU318817
6	206.4	99.2	633	1	MTU318814
7	206.4	99.2	633	1	MTU318816
8	206.4	99.2	635	1	MTU318821
9	206.4	99.2	970	6	150706
10	206.4	99.2	3534	6	AX111339
11	206.4	99.2	3853	1	MTU12205
12	206.4	99.2	5084	1	MSGRPOB
13	206.4	99.2	19352	1	AE006964
14	206.4	99.2	19770	1	MTU1376
15	147.2	70.8	3447	6	AR067447
16	147.2	70.8	3752	1	MSU24494
17	147.2	70.8	37617	1	MLB1790G
18	147.2	70.8	348950	1	MLEPRINT
19	98	47.1	3941	1	AF242549
20	88	42.3	3495	6	AX120631
21	88	42.3	328050	1	AP005275
22	88	42.3	349980	6	AX127144
23	80.4	38.7	32923	1	SCD82
24	45	21.6	15898	1	AX024393
25	44	21.6	15898	6	AX024286
26	44	21.2	7954	1	TA019223
27	43.4	20.9	134816	2	CNS07C9R
28	42.6	20.5	1416	1	CNS08CP9
29	42.6	20.2	139467	1	SRECCAGEN
30	41.8	20.1	34182	1	SC111
31	41.4	19.9	195859	14	AF281817
32	41.1	19.7	204050	1	AL646070
33	40.2	19.3	11103	1	AE004774
34	40.2	19.2	138792	2	AC113583
35	40	19.2	204050	1	AL646073
36	39.8	19.1	13431	1	AE005044
37	39.8	19.1	119972	2	AP004029
38	39.8	19.1	126038	8	AP000367
39	39.6	19.0	598	1	PRE316161
40	39.6	19.0	1271	1	SVU04837
41	39.6	19.0	1392	1	PSRECAO
42	39.6	19.0	3056	1	PPU70864
43	39.4	18.9	5894	3	ACMHC
44	39.4	18.9	44109	1	SCSFI
45	39.4	18.9	44109	1	SCSFI

ALIGNMENTS

RESULT 1
LOCUS MTU318818 610 bp DNA circular BCT 09-AUG-2002
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1415-97.
ACCESSION AJ318818
VERSION AJ318818.1 GI:22208412
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE 1
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE Molecular analysis of rifampin-resistant *Mycobacterium tuberculosis* isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 610)

AUTHORS Herrera, L.

JOURNAL Direct Submission

TITLE Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid 28220, SPAIN

FEATURES

source location/Qualifiers

1..610

/organism="Mycobacterium tuberculosis"

/isolate="1415-97"

/db_xref="taxon:1773"

1..610

/gene="rpoB"

<1..>610

/gene="rpoB"

/codon_start=1

/transl_table=1

/product="RNA polymerase beta subunit"

/protein_id="CAC87035.1"

/db_xref="GI:22208413"

/translation="LDIYKRLRPEPTRESAQTLLNLFPEKRYDLARVGRYVVK KLGLHVEPIITSSLTLEEDVATIEYLVLHSGOTTMTVPGVEVPEVETDIDHGNR RLRTVGLIQNIIVGMSRMERYRERMTQDEATTPOTLINIRPVVAALKKEFGTS QPSQFMGONNPLSLGLTHKRLSLALGPGSLSRERAGLEVDRVDP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 99.28; Score 206.4; DB 1; Length 610;

Best Local Similarity 99.58; Pred. No. 1.5e-33;

Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGAGAACGCGTACGACCTGGCCCGCGTGCCTATTAAGTCAACAAAGACTCG 60

DB 80 TCAAGAGAACGCGTACGACCTGGCCCGCGTGCCTATTAAGTCAACAAAGACTCG 139

QY 61 GCGTCATGTGGCGGACCCATCAGTCGTGACGCTGACCGAAGAAGCGTGGCCA 120

DB 140 GCGTCATGTGGCGGACCCATCAGTCGTGACGCTGACCGAAGAAGCGTGGCCA 199

QY 121 CCATCGAATATCTGTCGCGTGCAGAGGGTCAGACCAAGATGATCGTCCGGGGCGG 180

DB 200 CCATCGAATATCTGTCGCGTGCAGAGGGTCAGACCAAGATGATCGTCCGGGGCGG 259

QY 181 TCGAGGTGCCGCTGGAACCGACGACAT 208

DB 260 TCGAGGTGCCGCTGGAACCGACGACAT 287

RESULT 2

LOCUS MTU318819 610 bp DNA circular BCT 09-AUG-2002

DEFINITION *Mycobacterium tuberculosis* partial rpoB gene for RNA polymerase beta subunit, isolate 1417-97.

ACCESSION AJ318819

VERSION AJ318819.1 GI:22208414

KEYWORDS RNA polymerase beta subunit; rpoB gene.

SOURCE *Mycobacterium tuberculosis*.

ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE 1

AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.

JOURNAL Molecular analysis of rifampin-resistant *Mycobacterium tuberculosis* isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

REFERENCE 2 (bases 1 to 610)

AUTHORS Herrera, L.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid 28220, SPAIN

FEATURES

source location/Qualifiers

1..610

/organism="Mycobacterium tuberculosis"

/isolate="1417-97"

/db_xref="taxon:1773"

1..610

/gene="rpoB"

<1..>610

/gene="rpoB"

/codon_start=1

/transl_table=1

/product="RNA polymerase beta subunit"

/protein_id="CAC87035.1"

/db_xref="GI:22208413"

/translation="LDIYKRLRPEPTRESAQTLLNLFPEKRYDLARVGRYVVK KLGLHVEPIITSSLTLEEDVATIEYLVLHSGOTTMTVPGVEVPEVETDIDHGNR RLRTVGLIQNIIVGMSRMERYRERMTQDEATTPOTLINIRPVVAALKKEFGTS QPSQFMGONNPLSLGLTHKRLSLALGPGSLSRERAGLEVDRVDP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 99.28; Score 206.4; DB 1; Length 610;

Best Local Similarity 99.58; Pred. No. 1.5e-33;

Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGAGAACGCGTACGACCTGGCCCGCGTGCCTATTAAGTCAACAAAGACTCG 60

DB 80 TCAAGAGAACGCGTACGACCTGGCCCGCGTGCCTATTAAGTCAACAAAGACTCG 139

QY 61 GCGTCATGTGGCGGACCCATCAGTCGTGACGCTGACCGAAGAAGCGTGGCCA 120

DB 140 GCGTCATGTGGCGGACCCATCAGTCGTGACGCTGACCGAAGAAGCGTGGCCA 199

QY 121 CCATCGAATATCTGTCGCGTGCAGAGGGTCAGACCAAGATGATCGTCCGGGGCGG 180

DB 200 CCATCGAATATCTGTCGCGTGCAGAGGGTCAGACCAAGATGATCGTCCGGGGCGG 259

QY 181 TCGAGGTGCCGCTGGAACCGACGACAT 208

DB 260 TCGAGGTGCCGCTGGAACCGACGACAT 287

RESULT 3

LOCUS MTU318813 616 bp DNA circular BCT 09-AUG-2002

DEFINITION *Mycobacterium tuberculosis* partial rpoB gene for RNA polymerase beta subunit, isolate 1763-97.

ACCESSION AJ318813

VERSION AJ318813.1 GI:22208402

KEYWORDS RNA polymerase beta subunit; rpoB gene.

SOURCE *Mycobacterium tuberculosis*.

ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE 1

AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.

JOURNAL Molecular analysis of rifampin-resistant *Mycobacterium tuberculosis* isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

REFERENCE 2 (bases 1 to 616)

AUTHORS Herrera, L.

JOURNAL Direct Submission

TITLE Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid 28220, SPAIN

FEATURES

source location/Qualifiers

1..616

/organism="Mycobacterium tuberculosis"
/isolate="1763-97"
/db_xref="taxon:1773"
1..615
/gene="rpoB"
<1..>615
/gene="rpoB"

BASE COUNT 125 a 191 c 201 g 99 t
ORIGIN
Query Match 99.2%; Score 206.4; DB 1; Length 616;
Best Local Similarity 99.5%; Pred. No. 1.5e-33;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCCGCTACGACCTGGCCCGCTGCTGCTATTAAGTCAACAAGAGCTCG 60
DB 80 TCAAGGAGAGCCGCTACGACCTGGCCCGCTGCTGCTATTAAGTCAACAAGAGCTCG 139
QY 61 GCGTCATGTCGGCGAGCCCATCAGCTCGTCGACGCTGACCGAAGAAGCGTGGGCCA 120
DB 140 GCGTCATGTCGGCGAGCCCATCAGCTCGTCGACGCTGACCGAAGAAGCGTGGGCCA 199
QY 121 CCATCGAATATCTGTCGCCCTTGACAGAGGTCAACACGATGATCGTCCGGGCGCG 180
DB 200 CCATCGAATATCTGTCGCCCTTGACAGAGGTCAACACGATGATCGTCCGGGCGCG 259
QY 181 TCGAGGTGCCGCTGGAACCGACGACAT 208
DB 260 TCGAGGTGCCGCTGGAACCGACGACAT 287

RESULT 4
MTU318815 618 bp DNA circular BCT 09-AUG-2002
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 2540-97.
ACCESSION
AJ318815.1 GI:22208406
VERSION
RNA polymerase beta subunit; rpoB gene.
KEYWORDS
Mycobacterium tuberculosis.
SOURCE
Mycobacterium tuberculosis
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE
AUTHORS
Herrera, L., Jimenez, M. S. and Saez, J. A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
AUTHORS
Herrera, L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid, 28220, SPAIN
FEATURES
SOURCE
Location/Qualifiers
1..618
/organism="Mycobacterium tuberculosis"
/isolate="2540-97"
/db_xref="taxon:1773"
1..618
/gene="rpoB"
<1..>618

/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"
/translation="TDEALLDIYRLRPEPTKESAOILLLENLFKEKRYDLARVGR
YKVKRLGLHVEPTSSLTLEDVATIEYVRLHEQGTITPTPGVEVETDDID
HFGNRLRTVGLIONOIRVGSRRERVRMTTODVEATPTLNIIRPVAAIKE
FCTTSOLSPMDONNPLSGLTTHKRLALPGGLSRERAGLEVDRV"
BASE COUNT 124 a 190 c 207 g 97 t
ORIGIN

Query Match 99.2%; Score 206.4; DB 1; Length 616;
Best Local Similarity 99.5%; Pred. No. 1.5e-33;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCCGCTACGACCTGGCCCGCTGCTGCTATTAAGTCAACAAGAGCTCG 60
DB 95 TCAAGGAGAGCCGCTACGACCTGGCCCGCTGCTGCTATTAAGTCAACAAGAGCTCG 154
QY 61 GCGTCATGTCGGCGAGCCCATCAGCTCGTCGACGCTGACCGAAGAAGCGTGGGCCA 120
DB 155 GCGTCATGTCGGCGAGCCCATCAGCTCGTCGACGCTGACCGAAGAAGCGTGGGCCA 214
QY 121 CCATCGAATATCTGTCGCCCTTGACAGAGGTCAACACGATGATCGTCCGGGCGCG 180
DB 215 CCATCGAATATCTGTCGCCCTTGACAGAGGTCAACACGATGATCGTCCGGGCGCG 274
QY 181 TCGAGGTGCCGCTGGAACCGACGACAT 208
DB 275 TCGAGGTGCCGCTGGAACCGACGACAT 302

RESULT 5
MTU318817 618 bp DNA circular BCT 09-AUG-2002
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 2348-98.
ACCESSION
AJ318817.1 GI:22208410
VERSION
RNA polymerase beta subunit; rpoB gene.
KEYWORDS
Mycobacterium tuberculosis.
SOURCE
Mycobacterium tuberculosis
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE
AUTHORS
Herrera, L., Jimenez, M. S. and Saez, J. A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
AUTHORS
Herrera, L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid, 28220, SPAIN
FEATURES
SOURCE
Location/Qualifiers
1..618
/organism="Mycobacterium tuberculosis"
/isolate="2348-98"
/db_xref="taxon:1773"
1..618
/gene="rpoB"
<1..>618
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"

/translation="TDEALLDIYRKLRPGEPPTKESAOITLLENLFFKERYDLARV
YKVNKRLGLHVGEPITSTLTEDVAVATLEYLVRHLEGOTTTPVGGVEVETDID
HFGNRRLRTVGEILQNIQIRVGRSMRERVRRMTTODVEAITPQTLINIRPVAAIK
PFGTSQDSQSMGNPNLSGLTHKRRLSALGPGLSRRAGLEVRDV"
BASE COUNT 124 a 192 c 207 g 95 t
ORIGIN

Query Match 99.2%: Score 206.4; DB 1: Length 618;
Best Local Similarity 99.5%: Pred. No. 1.5e-33;
Matches 207: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TCAAGAGAACGCTACGACCTGGCCCGCTCGCTATAGTCAACAAAGACCTCG 60
|||||
Db 95 TCAAGAGAACGCTACGACCTGGCCCGCTCGCTATAGTCAACAAAGACCTCG 154
Oy 61 GGCTGCATGTGGGAGACCCATCAGCTGTCGACGCTGACCCGAAGAAGCTGTGGCCA 120
|||||
Db 155 GGCTGCATGTGGGAGACCCATCAGCTGTCGACGCTGACCCGAAGAAGCTGTGGCCA 214
Oy 121 CCATGAAATATCTGTGCTCGCTTCGACGAGGCTCAGACCAAGATGATGTTCCGGGGGGG 180
|||||
Db 215 CCATGAAATATCTGTGCTCGCTTCGACGAGGCTCAGACCAAGATGATGATCCGTTCCGGGGGGG 274
Oy 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
|||||
Db 275 TCGAGGTGCGCGGTGGAACCGACGACAT 302

RESULT 6
MTU318814 633 bp DNA circular BCT 09-AUG-2002
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1058-97.
ACCESSION
AJ318814
VERSION
AJ318814.1 GI:22208404
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE
AUTHORS
TITLE
Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
TITLE
Unpublished
REFERENCE
2 (bases 1 to 633)
AUTHORS
Herrera, L.
JOURNAL
TITLE
Direct Submission
SUBMITTED (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
LOCATION/Qualifiers
1. 633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1. 633
/gene="rpoB"
/codon_start=1
/transl_table=1
/product="RNA polymerase beta subunit"
/protein_id="CAC87033.1"
/db_xref="GI:22208405"
/translation="DNTVCTDEALLDIYRKLRPGEPPTKESAOITLLENLFFKERYDL
ARVGRYKNNKRLGLHVGEPITSTLTEDVAVATLEYLVRHLEGOTTTPVGGVEVETDID
IDHGNRRLRTVGEILQNIQIRVGRSMRERVRRMTTODVEAITPQTLINIRPVAAIK
EFEGTSQDSQSMGNPNLSGLTHKRRLSALGPGLSRRAGLEVRDV"

BASE COUNT 129 a 195 c 210 g 99 t
ORIGIN

Query Match 99.2%: Score 206.4; DB 1: Length 633;
Best Local Similarity 99.5%: Pred. No. 1.5e-33;
Matches 207: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TCAAGAGAACGCTACGACCTGGCCCGCTCGCTATAGTCAACAAAGACCTCG 60
|||||
Db 110 TCAAGAGAACGCTACGACCTGGCCCGCTCGCTATAGTCAACAAAGACCTCG 169
Oy 61 GGCTGCATGTGGGAGACCCATCAGCTGTCGACGCTGACCCGAAGAAGCTGTGGCCA 120
|||||
Db 170 GGCTGCATGTGGGAGACCCATCAGCTGTCGACGCTGACCCGAAGAAGCTGTGGCCA 229
Oy 121 CCATGAAATATCTGTGCTCGCTTCGACGAGGCTCAGACCAAGATGATGTTCCGGGGGGG 180
|||||
Db 230 CCATGAAATATCTGTGCTCGCTTCGACGAGGCTCAGACCAAGATGATGATCCGTTCCGGGGGGG 289
Oy 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
|||||
Db 290 TCGAGGTGCGCGGTGGAACCGACGACAT 317

RESULT 7
MTU318816 637 bp DNA circular BCT 09-AUG-2002
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1255-98.
ACCESSION
AJ318816
VERSION
AJ318816.1 GI:22208408
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE
AUTHORS
TITLE
Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
TITLE
Unpublished
REFERENCE
2 (bases 1 to 637)
AUTHORS
Herrera, L.
JOURNAL
TITLE
Direct Submission
SUBMITTED (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
LOCATION/Qualifiers
1. 637
/organism="Mycobacterium tuberculosis"
/isolate="1255-98"
/db_xref="taxon:1773"
1. 637
/gene="rpoB"
/codon_start=1
/transl_table=1
/product="RNA polymerase beta subunit"
/protein_id="CAC87033.1"
/db_xref="GI:22208409"
/translation="VGTDEALLDIYRKLRPGEPPTKESAOITLLENLFFKERYDLARV
GRYKNNKRLGLHVGEPITSTLTEDVAVATLEYLVRHLEGOTTTPVGGVEVETDID
IDHGNRRLRTVGEILQNIQIRVGRSMRERVRRMTTODVEAITPQTLINIRPVAAIK
EFEGTSQDSQSMGNPNLSGLTHKRRLSALGPGLSRRAGLEVRDVPSH"

BASE COUNT 128 a 198 c 210 g 101 t
ORIGIN

Query Match 99.2%: Score 206.4; DB 1: Length 637;
Best Local Similarity 99.5%: Pred. No. 1.5e-33;
Matches 207: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TCAAGAGAACGCTACGACCTGGCCCGCTCGCTATAGTCAACAAAGACCTCG 60

|||||
Db 101 TCAAGAGAAAGCGCTACAGACCTGCGCGGTGCTCGCTATTAAGGTCAAGAAAGCTCG 160
OY 61 GGCTGATGTCGGCGAGCCCATCAGCTGTGTGACGCTTACCGAAGAACAGCTCGTGGCCA 120
Db 161 GGCTGATGTCGGCGAGCCCATCAGCTGTGTGACGCTTACCGAAGAACAGCTCGTGGCCA 220
OY 121 CCATGCAATATCTGCTGCGCTTGCACGAGGGTACAGACCATGATGCTTCGGGGGGCGG 180
Db 221 CCATGCAATATCTGCTGCGCTTGCACGAGGGTACAGACCATGATGCTTCGGGGGGCGG 280
OY 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
Db 281 TCGAGGTGCGCGGTGGAACCGACGACAT 308
RESULT 8
LOCUS MTU318821 639 bp DNA circular BCT 09-AUG-2002
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1071-98.
ACCESSION AJ318821
VERSION AJ318821.1 GI:22208418
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE 1
AUTHORS Herrera, L., Jimenez, M. S. and Saez, J. A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 639)
AUTHORS Herrera, L.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
National Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid, 28220, SPAIN
FEATURES
source
1. .639
/organism="Mycobacterium tuberculosis"
/isolate="1071-98"
/db_xref="taxon:1773"
1. .639
/gene="rpoB"
<1. .>639
/gene="rpoB"
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87038.1"
/db_xref="GI:22208419"
/translation="RRTDALDLYRKLRRGPPPTKESQTLLENLFPEKRYDLAR
GRYVKKKGLAHGEPITSTLIEDVVAITIELVLHGGTMTVPGVGVETRD
IDHFGNRRLTVELLIONOIRVMSRMERVRERMTTOVEAITPTLLINIPAVAAI
KEPFGTSSLSQFVWONNPLSGILQKRRLSALRPGGLSRERAGLEVADVPHSHS"
BASE COUNT 126 a 202 c 212 g 99 t
ORIGIN
Query Match 99.2%; Score 206.4; DB 1; Length 639;
Best Local Similarity 99.5%; Pred. No. 1.5e-33;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TCAAGAGAAAGCGCTACAGACCTGCGCGGTGCTTAAAGTCAACAGACGCTCG 60
Db 101 TCAAGAGAAAGCGCTACAGACCTGCGCGGTGCTTAAAGTCAACAGACGCTCG 160
OY 61 GGCTGATGTCGGCGAGCCCATCAGCTGTGTGACGCTTACCGAAGAACAGCTCG 120
Db 161 GGCTGATGTCGGCGAGCCCATCAGCTGTGTGACGCTTACCGAAGAACAGCTCG 220

OY 121 CCATGCAATATCTGCTGCGCTTGCACGAGGGTACAGACCATGATGCTTCGGGGGGCGG 180
Db 221 CCATGCAATATCTGCTGCGCTTGCACGAGGGTACAGACCATGATGCTTCGGGGGGCGG 280
OY 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
Db 281 TCGAGGTGCGCGGTGGAACCGACGACAT 308
RESULT 9
LOCUS 150706 970 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 1 from patent US 5643723.
ACCESSION 150706
VERSION 150706.1 GI:2472409
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 970)
AUTHORS Persing, D. H., Hunt, J. J., Young, K. K. Y., Felmlie, T. A., Roberts, G. D.
and Whelan, A. Christian.
TITLE Detection of a genetic locus encoding resistance to rifampin in
mycobacterial cultures and in clinical specimens
JOURNAL Patent: US 5643723-A I 01-JUL-1997;
FEATURES
source
1. .970
/organism="unknown"
BASE COUNT 182 a 302 c 330 g 156 t
ORIGIN
Query Match 99.2%; Score 206.4; DB 6; Length 970;
Best Local Similarity 99.5%; Pred. No. 1.4e-33;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TCAAGAGAAAGCGCTACAGACCTGCGCGGTGCTGCTATTAAGTCAACAGAACTCG 60
Db 26 TCAAGAGAAAGCGCTACAGACCTGCGCGGTGCTGCTATTAAGTCAACAGAACTCG 85
OY 61 GGCTGATGTCGGCGAGCCCATCAGCTGTGTGACGCTTACCGAAGAACAGCTCGTGGCCA 120
Db 86 GGCTGATGTCGGCGAGCCCATCAGCTGTGTGACGCTTACCGAAGAACAGCTCGTGGCCA 145
OY 121 CCATGCAATATCTGCTGCGCTTGCACGAGGGTACAGACCATGATGCTTCGGGGGGCGG 180
Db 146 CCATGCAATATCTGCTGCGCTTGCACGAGGGTACAGACCATGATGATGATGCTTCGGGGGGCGG 205
OY 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
Db 206 TCGAGGTGCGCGGTGGAACCGACGACAT 233
RESULT 10
LOCUS AX111339 3534 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 2072 from Patent WO0123604.
ACCESSION AX111339
VERSION AX111339.1 GI:13927631
KEYWORDS
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacteriineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE 1 (bases 1 to 3534)
AUTHORS Bergeron, M. G., Boissinot, M., Huelstsky, A., m Nard, C., Ouellette, M.,
Picard, F. J. and Roy, P. H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 2072 05-APR-2001;
FEATURES Infectio Diagnostic (I.D.I.) INC. (CA)
Location/Qualifiers

source 1..3534
/organism="Mycobacterium tuberculosis"
/strain="Rv"
/db_xref="taxon:1773"

BASE COUNT 679 a 1081 c 1188 g 586 t

ORIGIN

Query Match 99.2%; Score 206.4; DB 6; Length 3534;
Best Local Similarity 99.5%; Pred. No. 1.3e-33;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCACGAGAACGCCCTACGACCTGGCCCGCGTGGCTATTAAGTCAACAAGAACCTCG 60
|||||
Db 902 TCACGAGAACGCCCTACGACCTGGCCCGCGTGGCTATTAAGTCAACAAGAACCTCG 961

OY 61 GCGTCATGTCGGCGAGCCCATCAGCTCGTCGACGCTGACCGAAGAACGTCGTGGCCA 120
|||||
Db 962 GCGTCATGTCGGCGAGCCCATCAGCTCGTCGACGCTGACCGAAGAACGTCGTGGCCA 1021

OY 121 CCATCGAATATCTGTCGCTTGCACAGAGGTCGACGACGATGATCTTCCGGCGCGC 180
|||||
Db 1022 CCATCGAATATCTGTCGCTTGCACAGAGGTCGACGACGATGATCTTCCGGCGCGC 1081

OY 181 TCGAGGTCCGCGTGGAAACCGACGACAT 208
|||||
Db 1082 TCGAGGTCCGCGTGGAAACCGACGACAT 1109

RESULT 11
MTU12205 3853 bp DNA linear BCT 02-MAR-2000
LOCUS
DEFINITION Mycobacterium tuberculosis H37Rv RNA-polymerase beta subunit (rpoB)
ACCESSION U12205
VERSION U12205.1 GI:515684
KEYWORDS
SOURCE
ORGANISM Mycobacterium tuberculosis.
Mycobacterium tuberculosis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacteriaceae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 3853)
Imboden, P., Toller, R., Marchesi, F., Telenti, A., Bodmer, T.,
Cole, S., Schopfer, K. and Burkart, T.
The rpoB gene of Mycobacterium tuberculosis
Unpublished
2 (bases 1 to 3853)
Imboden, P.
Direct Submission
Submitted (11-JUL-1994) Paul Imboden, Institute for Medical
Microbiology, University of Berne, Friedbuehlstrasse 51, Berne,
3010, Switzerland
Location/Qualifiers
1..3853
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
576..>3853
/gene="rpoB"
576..>3853
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA-polymerase beta subunit"
/protein_id="AAA20242.2"
/db_xref="GI:7144499"
/translation="MLEGCIADSRQSKTAASPSRPOSSNNNSVPGAPNRVSPAKL
REPLEVPLDVOYDSEBMLGSRPMRSAAERDQVDPVGLLEVLTELSTEDFSGS
MSLSFSDRPDVAAPDECKDKMTTAAPLFTVAEPINNNTGEIKSQTFMGDPFM
TEKGTFTIINGTERVVSQVRSQVPEDETIDKSTDKLSHVKVIPSQGALEFDVAK
RDVQVIRIDRKROPVIVLALKALMTSEQIVERGSEFIMRSTLEKNTQVDEALD
IYRLRGPPEPTKESATLLENLFKREKRYDLARVGRYKVKKKGGLHVGEPITSSTLT
EDVAVATIEYLRLHEGOTMTVPGGVEVPEPTDIDHFCRRRLRTYAGELONIRVG

MSMRVVERBMTTDOVEATTPOFLINIRPVNAIKKEPSTQSLSQPMONNPLSGLT
HKRLSLGEGCSREAGLEVPDYHRSRGRCPIETEPGPITGLIGLSIYARNP
FGTETPTRYKVVGVSDGIVTLADEEDHVVQANSPLIDAGKRVPEPVLRKAG
EVERVPSSEVDYMDVSPRQNSVATAMIPELEHDDANRALMGAMROQAVPLRSEAP
LVGKGMELRAAIDAATSRSQSESVIEBSADYITVHMDNGTRRTYMRFRARSNHTC
ANOCPIVDAGDRVACGVIADGCTDGEALGNLLVAIMPENHEDAIILSNRL
VEEDVLTSHIEEHEIDARDTKLGAETIDIDINISDEVADLDEGRIVRIGAEVSD
DILVGKVPKGTETETPEERLLAIGEKAREPRDSIKVPHRESKRVIGIRFSNED
EDELPRACVNELVRYVAAQKRISDGBKLGRHGNKVIGITLPEVDMPLADGTPVDI
ILNTHGVPRRMNIGQILLETHLGMCASHGKMWDAKVPMAARLPDELLAQPNATVS
TPVFDGAQEAELQGLLSCTLPNRDGVLYDADGKAMLFDRSGSEPPPYPTVGYVIM
KLHLVYDVKIHAIRSTGYSMTQOPLGKAKOFGQRFGEKMAQAVCAATVLOGLL
TKRS"

BASE COUNT 723 a 1173 c 1293 g 664 t

ORIGIN

Query Match 99.2%; Score 206.4; DB 1; Length 3853;
Best Local Similarity 99.5%; Pred. No. 1.3e-33;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TCACGAGAACGCCCTACGACCTGGCCCGCGTGGCTATTAAGTCAACAAGAACCTCG 60
|||||
Db 1477 TCACGAGAACGCCCTACGACCTGGCCCGCGTGGCTATTAAGTCAACAAGAACCTCG 1536

OY 61 GCGTCATGTCGGCGAGCCCATCAGCTCGTCGACGCTGACCGAAGAACGTCGTGGCCA 120
|||||
Db 1537 GCGTCATGTCGGCGAGCCCATCAGCTCGTCGACGCTGACCGAAGAACGTCGTGGCCA 1596

OY 121 CCATCGAATATCTGTCGCTTGCACAGAGGTCGACGACGATGATCTTCCGGCGCGC 180
|||||
Db 1597 CCATCGAATATCTGTCGCTTGCACAGAGGTCGACGACGATGATCTTCCGGCGCGC 1656

OY 181 TCGAGGTCCGCGTGGAAACCGACGACAT 208
|||||
Db 1657 TCGAGGTCCGCGTGGAAACCGACGACAT 1684

RESULT 12
MSGRPOB 5084 bp DNA linear BCT 13-SEP-1994
LOCUS
DEFINITION Mycobacterium tuberculosis RNA polymerase beta-subunit (rpoB)
ACCESSION L27989
VERSION L27989.1 GI:466333
KEYWORDS
SOURCE RNA polymerase beta-subunit; rpoB gene.
Mycobacterium tuberculosis (strain Rv) DNA.
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteridae; Actinobacteriaceae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 5084)
Miller, L.P., Crawford, J.T. and Shinick, T.M.
The rpoB gene of Mycobacterium tuberculosis
Antimicrob. Agents Chemother. 38 (4), 805-811 (1994)
MEDLINE 94304130
PUBMED 8031050
Location/Qualifiers
1..5084
/organism="Mycobacterium tuberculosis"
/strain="Rv"
/db_xref="taxon:1773"
1065..4598
/gene="rpoB"
1065..4598
/gene="rpoB"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="RNA polymerase beta-subunit"
/protein_id="AA21416.1"
/db_xref="GI:466334"
/translation="MLEGCIADSRQSKTAASPSRPOSSNNNSVPGAPNRVSPAKL

REPLEVGLLDVOTDSFEMLIGSPRMRESAERGDVNPVGCLEVLIELSPEDFSGS
MSLSFSDPRDDVAPYBECKDKMDTAPAEIPIAEFINNNGEIKQYFMGDFPMK
TEKGTFTINGTERVVSQVLRSPGVYFDEITDKSDKTLHSVKIIPSRGAMLEFDVAK
RDTVGRIIDRRKRPVTLKALGWTSEQIVVERGSEIMSTLEKNVTGDTALLD
IYRLKRGEPPTESAOTLENLFEKREKRYDLARGRYKKTGLGVEGPIITSSLT
EEDVATIEYLVRHLEGOTMTVPGVVEVPEOTDIOHFGRRIRJYGEILIONIRG
MSRMVERBERTODVEAITPOTLINIRPVVAIKERFETQSLSQMDONPLSGLT
HRRISALGCGLSREBRAGLEVDRVPHSHYRMCPIETPEBPNIGLIGSLSVARVP
RGPFIETPRKRVYDVSEIYLIADDEDRHVAQANSPIADAGRFEVPLVRKAG
EVEYPSSEVDMDVSPROMVSATAMIPLEHDANRALMGAMQOAVPLVSEAP
LVGTGAEMLKAAIDAATSSQSGVIEEVSADYITVMHNGRTYRMRKPARSHGTC
ANOCPIVADGDRVEAGVYIADGPTDDEMLGNLVAIMPGEHNVEDALITSNL
VEEDVATIEYLVRHLEGOTMTVPGVVEVPEOTDIOHFGRRIRJYGEILIONIRG
DILVGTPIKGETELTPEERILRAIFGKADEVDTSKPHGSGKVIGIRVFSRED
EDELPAQVNELRYRYVAOKRKISDGDLAGHNGKGVIGILPEVMDPLADGTPVDI
ILNTHGVPFRMNIGIOLLETHLGMCAHSGMKVDAKGVDMARLPDELLEAHANATIS
TVFDGAEBAELQGLSCTLPNRDGDVYVDADGKAMLPDGRSGEPFPYPTVYVMGI
KLHLHYDVKIHARSTGYSMTIQOPLGKAOFQGRGEMECMAQVGAAYTLOEL
TIKSDTVGRVYVEAIVKGENIPEPGIPESFKVYLKELQSLCLNVEVLSDDAATIEL
REGDEDELERAAANGILNLSNESASFEDLA"

gene
4641..5084
/gene="tpoc"
/note="tpoc"
4641..>5084
/gene="tpoc"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta'-subunit"
/protein_id="AAK4147.1"
/db_xref="GI:537608"
/translation="MLDVNFEDELIRIGLATAEADIRQMSYGEVKKPETINRYTLKPEKD
GLFCEKIFGPTDMECYGKKRVRFKGIICERCGVETRAKVRREMGHIELAAVYT
HIWFKGVPSSLGYLLDPLADEKIIYFAAYVITVSDEMRNEL"

CDS
1534 c 1691 g 890 t

BASE COUNT 969 a 1534 c 1691 g 890 t

ORIGIN

Query Match 99.2% Score 206.4 DB 1 Length 5084
Best local similarity 99.5% Pred. No. 1.3e-33
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TCACAGACAGAGCGCTACACACCTGGCCGCGGTACGCTAAGCTACACAGAGCTCG 60
1966 TCACAGACAGAGCGCTACACACCTGGCCGCGGTACGCTAAGCTACACAGAGCTCG 2025
61 GCCTCATGTGCGGAGCGCCATCACGTGTCAGACGTCACGAGAGAGAGCTCGCA 120
2026 GCCTCATGTGCGGAGCGCCATCACGTGTCAGACGTCACGAGAGAGAGCTCGCA 2085
121 CCATCAATATCTGCTCGGCTTGACAGAGGTCAGACGATGATGCTTCCGGCGGCG 180
2086 CCATCAATATCTGCTCGGCTTGACAGAGGTCAGACGATGATGCTTCCGGCGGCG 2145
181 TCGAGGTGCGGAGGAGAACGACGACAT 208
2146 TCGAGGTGCGGAGGAGAACGACGACAT 2173

RESULT 13

AE006964 19352 bp DNA linear BCT 27-APR-2001
LOCUS Mycobacterium tuberculosis CDC1551, section 50 of 280 of the
DEFINITION complete genome.
ACCESSION AE006964 AE000516
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
Mycobacterium tuberculosis CDC1551.
Mycobacterium tuberculosis CDC1551.
Bacteria: Firmicutes: Actinobacteria: Actinobacteridae:
Actinomycetales: Corynebacteriaceae: Mycobacteriaceae:
Mycobacterium: Mycobacterium tuberculosis complex.
1 (bases 1 to 19352)
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O.,
Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E.,
Kolony, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

Salzberg, S.L., Delcher, A., Uitterback, T., Weidman, J., Khouri, H.,
Gill, J., Mikula, A. and Bishai, W.
Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains
Unpublished
2 (bases 1 to 19352)
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O.,
Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E.,
Kolony, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M.,
Salzberg, S.L., Delcher, A., Uitterback, T., Weidman, J., Khouri, H.,
Gill, J., Mikula, A. and Bishai, W.
Direct Submission
Submitted (25-Apr-2001) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
Location/Qualifiers
1..19352
/organism="Mycobacterium tuberculosis CDC1551"
/strain="CDC1551"
/db_xref="taxon:83331"
/note="clinical strain"
163..3699
/gene="MT0695"
163..3699
/note="MT0695"
PID:149992; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA-directed RNA polymerase, beta subunit"
/protein_id="AAK4921.1"
/db_xref="GI:13880218"
/translation="MLEGCILADRSKSTAAASPPSPRSSSSNNSVPGANPRVFAKL
REPLEVGLLDVOTDSFEMLIGSPRMRESAERGDVNPVGCLEVLIELSPEDFSGS
MSLSFSDPRDDVAPYBECKDKMDTAPAEIPIAEFINNNGEIKQYFMGDFPMK
TEKGTFTINGTERVVSQVLRSPGVYFDEITDKSDKTLHSVKIIPSRGAMLEFDVAK
RDTVGRIIDRRKRPVTLKALGWTSEQIVVERGSEIMSTLEKNVTGDTALLD
IYRLKRGEPPTESAOTLENLFEKREKRYDLARGRYKKTGLGVEGPIITSSLT
EEDVATIEYLVRHLEGOTMTVPGVVEVPEOTDIOHFGRRIRJYGEILIONIRG
MSRMVERBERTODVEAITPOTLINIRPVVAIKERFETQSLSQMDONPLSGLT
HRRISALGCGLSREBRAGLEVDRVPHSHYRMCPIETPEBPNIGLIGSLSVARVP
RGPFIETPRKRVYDVSEIYLIADDEDRHVAQANSPIADAGRFEVPLVRKAG
EVEYPSSEVDMDVSPROMVSATAMIPLEHDANRALMGAMQOAVPLVSEAP
LVGTGAEMLKAAIDAATSSQSGVIEEVSADYITVMHNGRTYRMRKPARSHGTC
ANOCPIVADGDRVEAGVYIADGPTDDEMLGNLVAIMPGEHNVEDALITSNL
VEEDVATIEYLVRHLEGOTMTVPGVVEVPEOTDIOHFGRRIRJYGEILIONIRG
DILVGTPIKGETELTPEERILRAIFGKADEVDTSKPHGSGKVIGIRVFSRED
EDELPAQVNELRYRYVAOKRKISDGDLAGHNGKGVIGILPEVMDPLADGTPVDI
ILNTHGVPFRMNIGIOLLETHLGMCAHSGMKVDAKGVDMARLPDELLEAHANATIS
TVFDGAEBAELQGLSCTLPNRDGDVYVDADGKAMLPDGRSGEPFPYPTVYVMGI
KLHLHYDVKIHARSTGYSMTIQOPLGKAOFQGRGEMECMAQVGAAYTLOEL
TIKSDTVGRVYVEAIVKGENIPEPGIPESFKVYLKELQSLCLNVEVLSDDAATIEL
REGDEDELERAAANGILNLSNESASFEDLA"

gene
3744..7694
/gene="MT0696"
/note="MT0696"
/note="similar to SP:p37871; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA-directed RNA polymerase, beta-prime subunit"
/protein_id="AAK44922.1"
/db_xref="GI:13880219"
/translation="MLDVNFEDELIRIGLATAEADIRQMSYGEVKKPETINRYTLKPEKD
GLFCEKIFGPTDMECYGKKRVRFKGIICERCGVETRAKVRREMGHIELAAVYT
HIWFKGVPSSLGYLLDPLADEKIIYFAAYVITVSDEMRNEL"
AVEODGDESLRAQKLEADLAELABEBAADARVGRGERSKROQRDOAGREAS
LEDINSTPTKLAPOKLIYDENYRELVRITGCTFGAGAGSIOLEINFDIOEAS
LRVITNGCGKRLAKRLKLVYAAFGQSGNSPMQVLDVAVPIPELIRPVQDGR
FATSDLANDIRKVINNNKRLIDLGAPEIIVNNEKMLQESVATIDFNGRGRPVY
GPNRPLKLSLQKGRFONLGRVYSGRSYIVGPOKLKLOCGGLPKLMAL
LKPFPVKRLVDLNAONIKSAKRVREOROVMDVLEEVYAEHPLVLRAPTLTRHG
IOAFEPMLVEGKAIOHLPLVCAFNADPDGDMVNHPLSLSAQAEARILMLSSNLT

SPASGRPLAMPRLDMVTLGYLTTEVPEDTGEYOPASGDHPETGVYSSPAEIAIMADR
GVLVSRAKIVRLTOLRPLVEIEALFEGHSGWQGDAMMAETTLGVMNELILPGP
FVNKOHMKKVOAAIINDIAERYPMIYAQOTVDLXKACGPWMAKRSVYTSMADEVLP
RKKEIILHVEERADYKORQKALHNDENEALVEIMKATDEVOALREYDNDPN
IITIVBSGATGNTQRTTLGAKMGLVNTNKGFTIPRPVSSFEKGLTLEYEINTGA
RKGLADALTADSGYTLRLVDSQDIVREHDCOTERQIVVELAERAPDGLIIDP
YIETSAARTLGTDAVEAGNIVVERGODIDPEIDALLAAGITLOVYKMSVGLCATP
GVCATCGSRMAGTKLVDIGEAAGIIVAAOSIGEPQIOLMTRTFHOGVDEBITGCLPR
VOELFEARVGRKAPITADYTVRVRLEDGERFYKITTVPDGCGEVYDOKISKRORLY
FKHEHDSERYLSGDHVEVGOOLMECSADPHVLRQGREVOTHLVREVOEYPRAG
VSTHDKHIEYIVOMLRVYITIDSGSTELPQSLIDRAEFEAENRRVAVAGEEPAGR
PVMGITKASLATDSMLSAASFQETTRVLTDAAINCRSDKLNGLKENVIGIKLPGT
GINKRRIAVQPEFEARAAYTTIPSYEDDYSPDFGAATGAVPILDYGYSDVR"
/note="MT0697"
/complement(7691. .8065)
/gene="MT0697"
/complement(7691. .8065)
/gene="MT0697"
/note="identified by glimmer2: putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44923.1"
/db_xref="GI:13880220"
/translation="MPSAAATINPCNANASAMERSGLLECVAAGLDEOPGFEFTADKL
NPDGSSRRVPRROADGIAITHVERGGQROSGGQGVVPRRHGFPALAMODRLIH
GEOTQNRIDQAEVRFVRCVSP"
/complement(8058. .9972)
/gene="MT0698"
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact; identified by
glimmer2: putative; conserved hypothetical protein",
authentic frameshift"
10167. .10925
/gene="MT0699"
10167. .10925
/gene="MT0699"
/note="identified by match to PRM protein family HM
PF01261"
/codon_start=1
/transl_table=11
/product="AP endonuclease, family 2"
/protein_id="AAK44924.1"
/db_xref="GI:13880221"
/translation="MLIGSHVSPDPIAAAEAGADVQVPIFGNPSKAKKPRDDAA
ALKATIPPIYHAPYLINLASANNRYRIPSRITLOETCAADIGAAAVIHGGVAD
DNDIDKGRMRKALDRLETEVPYVLENTAGSDHMAKRFDTIARLMDVIDTGGFC
LDCTHTWAGEALTDADVRIKAITGRIDLPHGNDSDRDEAGSGRDRHAMLGSQIDPDL
LVAAVKAAGAPVICTADQGRKDDIAFLKERTGS"
10957. .11799
/gene="MT0700"
10957. .11799
/gene="MT0700"
/note="similar to GB:U00012 PID:466863; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hydrolase/esterase, putative"
/protein_id="AAK44925.1"
/db_xref="GI:13880222"
/translation="MLRVAAILLAVALAFACSGSTRLAAGFGNGNSVHTLDVAGR
SYRLKPVGLPSSAPLVVMLHGFSGAKQASRGYDELADSEKFLVAVPGYHRAN
ANGGCCGRPARREGVDIGFRAVAVDIAANNVSDIPARVYTGSMNCAIMSYTACNT
SIFAIIGVSGTOLDPCOSPRPVSYIHHGADPLVHRHGGPGAFATIDPPYDILN
AFWRVNRGALDITTBGPVTTSGATCADNRRVYLLIYDDAGHRMPSFATOTLWRFPA
AHRP"
11859. .13487
/gene="MT0701"
11859. .13487
/gene="MT0701"
/note="similar to SP:P33224 GB:L20915 PID:457172
PID:457174 PID:537028; identified by sequence similarity;
putative"
/codon_start=1

/transl_table=11
/product="acyl-CoA dehydrogenase, putative"
/protein_id="AAK44926.1"
/db_xref="GI:13880223"
/translation="MSDRHVYVNVQNPPLNENPASPVLIALLOEGGONCLDEVNEY
GATSSCAQNRKGLADNRRLIHTHDATIGRVDEVEDPYVHELMRLATIGHMAAP
WADDPGHVHVAARATSVTVPEGHICPISMTYAVPALRYNSLELAAYEPLTSREY
DEPLPATTKAGITAGMSMTKEKOGSDVRACTTOATPADGYSLTGHIKFTSAPKD
IFLVQAAPDGLSCFLPRVLDPDGRNMFQORIKLGNHANSSEVEYDGAAMLY
GEEGVPITIEMLVTLRDLCAISATSMRTGLRAVHAQHRKAFGAYLIDPILMR
VLADLAVEAEATIVAMAGATNANVGNTEFALLRIGLAAKHYKCRSTHAAE
ALEELGNGYVEDSGMPRLYREAPLAKMIGESSVNSALDTIRAAATPRACEVLEDEL
ARSAGQDRPLDGHVRLRPQGLDITIGYRARKIAEDICLALDGLSLVRHGHAVAEA
FLATRLSGQWGAAYGTMPAGIDLAPILTERALVKG"
13498. .14436
/gene="MT0702"
13498. .14436
/gene="MT0702"
/note="similar to GP:3885480; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="enoyl-CoA hydratase/isomerase family protein"
/protein_id="AAK44927.1"
/db_xref="GI:13880224"
/translation="MTHAIRPVDFDNLKTMTEYVTRGIARITENPEKNAIADTPL
ELSLVERADLDPEGVHVLIVSGRGEFCAGDLSAVAESSSTGGGAYOGYLDCKT
QAVNHPQWPMIDYQMSRFRWGRFASLHNAKPIYVVIHGCGVAGTDLIHLNDQ
VIAADAKIGYPTRWGVPAAGLMAHRRKARLLFTFGDITGQAAEMGLAVBA
PEPADDERTELRIARIALPVLNOLIVKALINSALLOQGVATSRVSTYFDGAHRT
PEGHAFVADAVEHGFRAVRRRDEPFIDYGRASRV"
14439. .15161
/gene="MT0703"
14439. .15161
/gene="MT0703"
/note="identified by glimmer2: putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"

Query Match 99.2%; Score 206.4; DB 1; Length 19352;
Best Local Similarity 99.5%; Pred. No. 1.2e-33;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TCAAGAGAGAGCGGTACACCTGGCCGCGTCCGTATTAAGTCAACAAGACCTCG 60
|||||
Db 1064 TCAAGAGAGAGCGGTACACCTGGCCGCGTCCGTATTAAGTCAACAAGACCTCG 60
QY 61 GGCTGATGTCGGCGAGCCATCACTGCTGACGCTGACCGAAGAAGACGTCGGCCA 120
|||||
Db 1124 GGCTGATGTCGGCGAGCCATCACTGCTGACGCTGACCGAAGAAGACGTCGGCCA 1183
QY 121 CCATCGAATATCTGTCGCTTCGACGAGGCTCAACACCATGATCTCCGGCGCGC 180
|||||
Db 1184 CCATCGAATATCTGTCGCTTCGACGAGGCTCAACACCATGATCTCCGGCGCGC 1243
QY 181 TCGAGGTCCGCGTGAACACGACGACAT 208
|||||
Db 1244 TCGAGGTCCGCGTGAACACGACGACAT 1271

RESULT 14
LOCUS MT01376 19770 bp DNA linear BCT 03-ANG-2001
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 32/162.
ACCESSION 295972 AL123456
VERSION 295972.1 GI:3261790
KEYWORDS
SOURCE
ORGANISM
Mycobacterium tuberculosis H37Rv.
Mycobacterium tuberculosis H37Rv.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE 1 (bases 1 to 19770)
 AUTHORS Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry III, C.E., Tekala, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Fellwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moulé, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A., Rajadaram, M.A., Rogers, J., Ruter, S., Seeger, K., Skelton, S., Squares, S., Squires, R., Sulston, J.E., Taylor, K., Whitehead, S., and Barrall, B.G.
 TITLE Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence
 JOURNAL Nature 393 (6685), 537-544 (1998)
 MEDLINE 98295987
 PUBMED 9634230
 REFERENCE 2 (bases 1 to 19770)
 AUTHORS Parkhill, J.
 TITLE Direct Submission
 JOURNAL Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
 COMMENT On Jun 27, 1998 this sequence version replaced gi:2143285.
 Notes:
 Details of *M. tuberculosis* sequencing at the Sanger Centre are available on the World Wide Web.
 (URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have been renumbered from the original cosmid submissions but the old gene designations are in brackets after the new gene numbers.
 Gene prediction was based on a Hidden Markov Model of 78 genes implemented in Reparse (Krogh) supplemented with visual inspection of positional base preference in codons, especially where there is an increase in the observed/expected third position G + C.
 CAUTION: In some cases we may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtg, or ttg) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.
 Location/Qualifiers
 source 1..19770
 /organism="Mycobacterium tuberculosis H37Rv"
 /strain="H37Rv"
 /db_xref="taxon:83332"
 /clone="1376"
 57..61
 /note="agga, possible rbs upstream of Rv0654"
 68..1573
 /gene="Rv0654"
 68..1573
 /note="Rv0654", (MTC1376.22), len: 501, unknown, FASTA score: 0.53153 LIGNOSTILBENE-ALPHA, BETA-DIOXYGENASE (485 aa) Opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity in 523 aa overlap). Also similar to *M. tuberculosis* protein MTCY21C12.07c (29.5% identity in 522 aa overlap).
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein Rv0654"
 /protein_id="CAB09380.1"
 /db_xref="GI:2143307"
 /db_xref="SPTREMBL:006785"
 /translation="MTTAAQASQNPYLGFLAPVSTEVATDLPVTGRIPEHLDRYLRNPNVAEVDPAATYVPTGDMAGVAVLRDQKAWYNNRWRTPAVCAALGEISARPHRTGIIIGGFRNTVLTAGRTALVEAGVAVNELTDELDTGVCDFDGLHGTYAHFQRDHTLHAIVSYFARGHRYQSVIGDGHARRVDIEVAGSPMHSFSLTDNVVYIYDLPVTFDPMQVAVSPRMQLQRPALVYQSVLGRVAPDPIAALGNMOGHSRLPYANPSTYPARVGVPRGNGEDVWMPDIEPCYVHPLNAYSECRRNGAEVLYVDYRYSRMEDRDRGSGDSPSLDWTINLATGAVTACRRDRAQOEPRINERTIVGSPHRYFATVIGEGFLVAGCAALSTPIKDCDCTGSSVTAASLDPLLIGEMVFPVPSRAEDDGLMGCGHRRDRDGLLLDAQTLESIAIVHLPRVPMGFHGNMAPTT"
 1585..2664
 /gene="Rv0654"

CDS
 /gene="Rv0655"
 1585..2664
 /gene="Rv0655"
 /note="Rv0655", (MTC1376.21), len: 359, abc transporter, FASTA score: TRBF_ECOLI P45393 hypothetical abc transporter atp-binding (269 aa) Opt: 644 z-score: 721.8 E(): 3.4e-33 (38.5% identity in 244 aa overlap); contains PS00017 ATP/GTP-binding site motif A, PS00211 ABC transporters family signature, highly similar to *M. leprae* MTL MYCUE P30769 possible ribonucleotide transport atp- in 335 aa overlap). Also similar to many others.
 tuberculosis ABC transporters eg. MTCY253.24 (33.6% identity in 241 aa overlap).
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein Rv0655"
 /protein_id="CAB09379.1"
 /db_xref="GI:2143306"
 /db_xref="SWISS-PROT:006784"
 /translation="MRYSDSYHTTGRMQPRASTEGFPMGVSIENVGLTKSGSSRIE DVTLLP:GGEVSVLLGPGTGSVFLKSLGLRPERGSIITDSTDIICSAEKLFI RTLFGVLPRDGLGSMNLVNTAPFPIREHKKESSIRIYVEKLAIVGLGDEKFI PGEISGHRKAGARALVLDPPQIILDEPDGSLDPVRYVLSQLMDINAQIDATIL IYTHNINARTVPPDMGMLEFKHLVMTGPPREVLITDEPVROFLNCRIRGIPGMSE KDEATWAEQALLDAGHHAGVEIEGVPOISATPMPERKAVARQARVREMLHTL PKKAVALIDDI:EGTHKVAHEIQ"
 1762..1785
 /gene="Rv0656"
 /note="PS00017 ATP/GTP-binding site motif A"
 2074..2118
 /gene="Rv0655"
 /note="PS00211 ABC transporters family signature"
 complement(3052..3435)
 /gene="Rv0656C"
 complement(3052..3435)
 /gene="Rv0656C"
 /note="Rv0656C", (MTC1376.20, unknown), len: 127 aa
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein Rv0656C"
 /protein_id="CAB09378.1"
 /db_xref="GI:2143305"
 /db_xref="SPTREMBL:006783"
 /translation="MAAATTTGHRGLLELAQRAVAGSCEPORAEPCRSARNADEFDQ MSRMGDIYDVPVPSKSVWRMIDSAQRIELARAAGVAGLSVYDILLICDTAARGLVYIH DDADYELAEHLDPIDIRVRYSAD"
 complement(3530..3685)
 /gene="Rv0657C"
 complement(3530..3685)
 /gene="Rv0657C"
 /note="Rv0657C", (MTC1376.19), unknown, len: 51 aa; similar to several other *M. tuberculosis* hypothetical proteins eg. YW08_MYCTU Q10848 hypothetical 8.9 kd protein cv39.08c (80 aa), FASTA scores: Opt: 107 z-score: 182.3 E(): 0.0039, 45.8% identity in 48 aa overlap. Also similar to MTCY48_5 and MTCY09581SC4H8.7 Streptomyces coelicolor cosmid 4H8 (66 aa), 41.0% identity in 39 aa overlap.
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein Rv0657C"
 /protein_id="CAB09377.1"
 /db_xref="GI:2143304"
 /db_xref="SPTREMBL:006782"
 /translation="MSVTQDLDLDELADVMRIAAVHTKKAENVLAMDYERFRIIE ALARSR"
 complement(3691..3696)
 /note="possible RBS upstream of Rv0657C"
 complement(3761..4477)
 /gene="Rv0658C"
 complement(3761..4477)
 /gene="Rv0658C"
 /note="Rv0658C", (MTC1376.18), len: 238, unknown,

gene

probablemembrane protein, similar to YPRB_ECOLI_P33774
hypothetical24.3 kd protein (urf 1) (217 aa), fasta
scores: opt: 174 z-score: 215.8 E(): 5.3e-05, (25.6%
identity in 223 aa overlap). Also similar to MTCY359.10
(28.7% identity in 178 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0658c"
/protein_id="CA809376.1"
/db_xref="GI:2143303"
/db_xref="SPTREMBL:O06781"
/translation="MEAGRADYVAPSHRWGLGAPLVVELVFLVASTSLAVVLTGHGPV
SAGVLALALAPYVAAAGLAILITRLRGLRSMNGLGLMFGGGLVLT
IPASLYTAIVGPEANSVAIRIEGGRAPALVPLVYVVAIDCEIITYRGLMG
AVDRMGRAALVYTVTFALAHLEPARAPLVVAIPALARFISGGLASIVTHOV
TNLPGLVILLGTTGALISLP"
complement(4480..4483)
/note="possible RBS upstream of Rv0658c"
complement(4753..5061)
/gene="Rv0659c"
complement(4753..5061)
/gene="Rv0659c"
/note="Rv0659c, (MTCI376.17), len: 102; unknown, similar
to YW28_MCTU_Q10867 hypothetical 12.3 kd protein cy39.28
(114 aa), fasta scores: opt: 144 z-score: 213.2 E():
7.3e-05, 30.8% identity in 107 aa overlap. Also similar to
MTCY09F9.22 (32.7% identity in 101 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0659c"
/protein_id="CA809387.1"
/db_xref="GI:2143302"
/db_xref="SPTREMBL:O06780"
/translation="MRGELMFAATPGGDRPLYLTPRPVADRIGAVVYVLTTRRG
LVSELELTAVENRVPDVCVNFNDNIHTLPRFAFRIRIRLSPALHENCOTLRASGCG
"
complement(5048..5293)
/gene="Rv0660c"
complement(5048..5293)
/gene="Rv0660c"
/note="Rv0660c, (MTCI376.16), len: 81; some similarity to
IAF016485.130 Halobacterium sp: NRC-1 plasm (100 aa),
32.4% identity in 74 aa overlap"
/codon_start=1
Query Match 99.2% Score 206.4; DB 1; Length 19770;
Best Local Similarity 99.5%; Pred. No. 1.2e-33;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTGCGCCCGGTGGTGCCTATTAAGTCAACAAGAGCTCG 60
|||||
Db 10758 TCAAGGAGAGCGCTACGACCTGCGCCCGGTGGTGCCTATTAAGTCAACAAGAGCTCG 10817
QY 61 GGCTGATGTGCGGAGCCCATACGCTGACCGCTGACCGAAGAGACGTCGTGGCCA 120
|||||
Db 10818 GGCTGATGTGCGGAGCCCATACGCTGACCGCTGACCGAAGAGACGTCGTGGCCA 10877
QY 121 CCATCGAATATCTGTCCTGCTGCACGAGGTCACACCAAGATGATCGTCCGGCGGCG 180
|||||
Db 10878 CCATCGAATATCTGTCCTGCTGCACGAGGTCACACCAAGATGATCGTCCGGCGGCG 10937
QY 181 TCGAGGTCCGCGTGAACCGACGACAT 208
|||||
Db 10938 TCGAGGTCCGCGTGAACCGACGACAT 10965
RESULT 15
AR067447 3447 bp DNA linear PAT 29-SEP-1999
LOCUS Sequence 57 from patent US 5851763.
DEFINITION AR067447
ACCESSION AR067447
VERSION AR067447.1 GI:5998669
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 3447)
AUTHORS Hymn B., Cole S., Young D., Zhang Y., Honore N., Telenti A. and
Bodmer J.
TITLE Rapid detection of antibiotic resistance in mycobacterium
tuberculosis
JOURNAL Patent: US 5851763-A 57 22-DEC-1998;
FEATURES Location/Qualifiers
source 1..3447
BASE COUNT 687 a 965 c 1139 g 656 t
ORIGIN
Query Match 70.8% Score 147.2; DB 6; Length 3447;
Best Local Similarity 81.7%; Pred. No. 3e-21;
Matches 170; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTGCGCCCGGTGGTGCCTATTAAGTCAACAAGAGCTCG 60
|||||
Db 809 TCAAGGAGAGAGCGCTACGACCTGCGCCCGGTGGTGCCTATTAAGTCAACAAGAGCTCG 868
QY 61 GGCTGATGTGCGGAGCCCATACGCTGACCGCTGACCGAAGAGACGTCGTGGCCA 120
|||||
Db 869 GGCTGATGTGCGGAGCCCATACGCTGACCGCTGACCGAAGAGATGTCGTGGCCA 928
QY 121 CCATCGAATATCTGTCCTGCTGCACGAGGTCACACCAAGATGATCGTCCGGCGGCG 180
|||||
Db 929 CCATAGAGTACCTGCTGCTGCTGCATGAGGTCACGATGACATGACTGCCAGGTGGCG 988
QY 181 TCGAGGTCCGCGTGAACCGACGACAT 208
|||||
Db 989 TAGAAGTCCAGTGAACCGACGATAT 1016

Search completed: November 13, 2002, 01:27:59
Job time : 660.723 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 ; Search time 1141.13 Seconds

(without alignments)
3037.202 Million cell updates/sec

Title: US-09-697-123b-21

Perfect score: 214
Sequence: 1 tcaagagagaagcgtacgat.....ccggtgagatgtgacacacat 214

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estlum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estcl:*
10: gb_estt2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estlum:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fut:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	137.6	64.3	1282	9 A1770311	A1770311 42 Mycoba
C 2	47.4	22.1	925	17 CNS0091P	AL053013 Drosophi
C 3	46	21.5	417	14 BM85094	BM85094 mgap008xj
C 4	44.6	20.8	569	17 BM897054	BM897054 3526_1_6
5	44	20.6	585	13 BM489075	BM489075 pgm2n.pk0
6	44	20.6	635	13 B1065677	B1065677 pgtfn.pk0

7	44	20.6	666	13 BM486272	BM486272 pgm1c.pk0
8	44	20.6	669	13 BM485879	BM485879 pgm1c.pk0
9	44	20.6	672	13 B1716865	B1716865 103101380
10	44	20.6	682	13 BM485533	BM485533 pgm1n.pk0
11	44	20.6	783	12 BF864011	BF864011 963048D05
12	43.8	20.5	568	13 B1995346	B1995346 1031026E0
13	43.4	20.3	710	12 BG321153	BG321153 zmo4_05f0
14	43.2	20.2	1856	11 A1109418	A1109418 zea mays
15	43	20.1	558	12 BE754521	BE754521 zea mays
16	42.8	20.0	275	14 R6542	R6542 RABEST103T
17	42.6	19.9	455	10 AV635603	AV635603 AV635603
C 18	42.6	19.9	523	9 A1856228	A1856228 sb39e01.x
19	42.4	19.8	487	14 B0824742	B0824742 1030120H0
20	42.4	19.8	923	17 CNS03YB3	AL262200 tetrarodon
C 21	42.2	19.7	553	9 A178937	A178937 zp39c04.s
22	42	19.6	370	12 BG159643	BG159643 OV2_7_C05
23	42	19.6	373	10 BM564128	BM564128 LG1_282_E
24	42	19.6	477	14 W17149	W17149 zb13b04.f1
25	42	19.6	593	10 BE358685	BE358685 DC1_31_D0
C 26	42	19.6	710	13 BG933256	BG933256 W81_92_B1
27	42	19.6	2598	11 AY103647	AY103647 zea mays
28	41.6	19.4	373	10 BE442717	BE442717 WHE1105_C
29	41.6	19.4	396	10 BE429399	BE429399 MTD017.F0
30	41.6	19.4	515	12 BF177122	BF177122 EM1_2_A10
31	41.6	19.4	587	14 B0743581	B0743581 WHE4105_F
32	41.6	19.4	636	14 B0294664	B0294664 WHE2853_A
33	41.4	19.3	234	14 R86588	R86588 RABEST163T
34	41.4	19.3	235	14 R86566	R86566 RABEST140T
35	41.4	19.3	265	14 E27471	F27471 HSPD15421.H
36	41.4	19.3	318	14 F00050	F00050 HSB09E061.S
37	41.4	19.3	331	14 R86510	R86510 RABEST202T
38	41.4	19.3	369	14 F36189	F36189 HSPD3571.H
39	41.4	19.3	376	14 F21475	F21475 HSPD06020.H
40	41.4	19.3	389	14 F25575	F25575 HSPD12678.H
41	41.4	19.3	392	9 AA180106	AA180106 zps4907.f
42	41.4	19.3	397	9 AA112975	AA112975 zn59c09.f
43	41.4	19.3	475	12 BF221644	BF221644 7059h10.x
C 44	41.4	19.3	502	9 AA211560	AA211560 zn55f10.f
45	41.4	19.3	502	14 BM698942	BM698942 UI-E-DX1-

ALIGNMENTS

RESULT 1
A1770311/c
LOCUS 1282 bp mRNA linear EST 24-JAN-2000
DEFINITION 42 Mycobacterium anaerobic stationary phase library Mycobacterium
smegmatis cDNA, mRNA sequence.
ACCESSION A1770311
VERSION A1770311.1 GI:6742680
KEYWORDS EST.
SOURCE Mycobacterium smegmatis.
ORGANISM Mycobacterium smegmatis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium.

REFERENCE 1 (bases 1 to 1282)
Munugasu-Oel B., Tay A. and Dick T.
Upregulation of stress response genes and ABC transporters in
anaerobic stationary-phase Mycobacterium smegmatis
Mol. Gen. Genet. 262 (4-5), 677-682 (1999)
JOURNAL MEDLINE
COMMENT Contact: Munugasu-Oel B.
Mycobacterium Laboratory
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore 117609, Republic of Singapore
Tel: 65 874 3011
Fax: 65 779 1117
Email: mcbomelmbc.nus.edu.sg
Insert length: 1282 Std Error: 0.00
Seq primer: T3 Forward; T7 Backward.
Location/Qualifiers
1..1282

FEATURES
source


```

/organism="Zea mays"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/clone_idb="3526" RescuedMu Grid K"
/clssue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: leaf; Vector: RescuedMu (engineered from
phluescript backbone); site_1: BamHI; Site_2: BglII;
RescuedMu is a 4.9 kb, modified maize Mu transposon

```

```

"job_xref": "taxon:9031",
"clone": "pgm2n.pk009.n11",
"clone_11b": "Normalized Chicken Breast Muscle, Leg Muscle,
and Epiphyseal Growth Plate cDNA Library (pgm2n)",
"/sex": "Male and Female",
"/tissue_type": "Breast muscle, leg muscle and epiphyseal
growth plate",
"/dev_stage": "Breast.leg:Embryo(d19).post-hatch1d,1,3,5,7,9,
11 weeks:Growth plate1d,7d,14d post-hatch",
"/lab_host": "E. coli EMP108",
"/note": "Vector: pCMVSPORT6; Library made from equivalent

```


LOCUS	BM485879	669 bp	mRNA	linear	EST 07-FEB-2002
DEFINITION	pgm1c.pk002.a9 Primary Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm1c) Gallus gallus cDNA clone pgm1c.pk002.a9 5' similar to sp1p025881PCS.CHICK TROPONIN C, SKELETAL MUSCLE p1r1TPCHS troponin C, skeletal muscle - chicken, mRNA sequence.				
ACCESSION	BM485879				
VERSION	BM485879.1 GI:18606809				
KEYWORDS	EST.				
SOURCE	chicken.				
ORGANISM	Gallus gallus				
REFERENCE	Archaeopteryx, Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Phasianidae, Gallus.				
AUTHORS	1 (bases 1 to 669)				
TITLE	Cogburn, L.A. and Monsonego-Ornan, E.				
JOURNAL	ESTs from Primary Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome Project Unpublished (2002)				
COMMENT	Contact: Larry A. Cogburn University of Delaware Townsend Hall, Newark, DE 19717, USA Tel: 302-831-1335 Fax: 302-831-2822 Email: cogburn@udel.edu, www.chickest.udel.edu.				
FEATURES	Location/Qualifiers				
SOURCE	1..669				
	/organism="Gallus gallus"				
	/strain="Commercial broiler and Ottawa Res. Centre strains 90 & 21"				
	/db_xref="taxon:9031"				
	/clone.lib="pk002.a9"				
	/clone.lib="Primary Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgm1c)"				
	/sex="Male and Female"				
	/tissue="type="Breast muscle, leg muscle and epiphyseal growth plate"				
	/dev_stage="Breast, leg: Embryo(d19):post-hatch(1d,1.3,5,7,9,11 weeks):growth plate(1d,7d,14d post-hatch)"				
	/lab_host="E. coli EMDH10B"				
	/note="Vector: pCMVSPORT6: Library made from equivalent pools of total RNA isolated from each tissue (embryonic muscle 33.3%, juvenile muscle 33.3%, and epiphyseal growth plate 33.3% of the final RNA pool). Single pass sequencing from 5'-end"				
BASE COUNT	152 a	172 c	230 g	112 t	3 others
ORIGIN					
Query Match	20.6%	Score 44;	DB 13;	Length 669;	
Best Local Similarity	53.5%	Pred. No. 6.2;			
Matches	92;	Conservative 0;	Mismatches 80;	Indels 0;	Gaps 0;
OY	41	CAAGGTGAACAAAGAGCTGGGCTGGGGGAGCACCATCGCTGAGTGGACACCACCAC	100		
Db	169	CACGCTGATGAGAGTGCCTGGGGCCAGAACCCACCAAGAGAGAGCTGGATGCCATCATTCGA	228		
OY	101	CCTCAGCAGAGAGAGAGCTGTGCCACCACCATGAGTACCTGTGTGGCCCTGCACAGAGGCCA	160		
Db	229	GGAGGTGACACAGAGTGTGCACGCGGCACCATCATGACATTCAGATGATGATGATGATGAT	288		
OY	161	GACACAGATGACCGCCCGCGGGCGGTGAGAGTGGCGGTGATGGAGCGAC	212		
Db	289	GCGCCAGATGAAAGAGACGCGCAAGGCGCAACTGTGAGAGAGAGCTGGCGCAC	340		
RESULT 9	672 bp				
LOCUS	BM485879	672 bp	mRNA	linear	EST 19-SEP-2001
DEFINITION	1031013B08.v1 C. reinhardtii CC-1690, Stress II (normalized), Lambda zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.				
ACCESSION	B1716865				

VERSION	BT1716865.1	GI:15692560
KEYWORDS	EST	
SOURCE	Chlamydomonas reinhardtii.	
ORGANISM	Chlamydomonas reinhardtii	
	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;	
	Chlamydomonadaceae; Chlamydomonas.	
REFERENCE	1 (bases 1 to 672)	
AUTHORS	Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre	
	,P., McDermott,J.P., Strager,J., Silflow,C. and Stern,D.	
TITLE	Analyses of the Chlamydomonas reinhardtii Genome: A Model,	
	Uncellular System for Analyzing Gene Function and Regulation in	
	Vascular Plants. Project: 1031	
JOURNAL	Unpublished (2001)	
COMMENT	Contact: Charles Hauser	
	DCMB Box 91000	
	Duke University	
	Durham, NC 27708-1003	
	Tel.: 919 613 8159	
	Fax: 919 613 8177	
	Email: chauser@duke.edu.	
FEATURES		
source	Location/Qualifiers	
	1..672	

FEATURES	SOURCE
Location/Qualifiers	1. 672
	/organism="Chlamydomonas reinhardtii"
	/strain="CC-1690 wild type mt+ 21gr"
	/db_xref="taxon:3055"
	/clone_lib="C. reinhardtii CC-1690, Stress II (normalized)
	/lambda Zap II"
	/note="vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; Stress condition II library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH4+ - containing) and shifted to TAP - NO3- (24hrs); H2 production conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with Exsistist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."
BASE COUNT	131 a 207 c 224 g 110 t
ORIGIN	
Query Match	20.6%; Score 44; DB 13; Length 672;
Best Local Similarity	52.8%; Pred. No. 6.2;
Matches	95; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
OY	35 TCGGACAAAGGTGAACAAGAGTGGCGCTGGGGCGGCACCAATCCGGCTCAGGTGACAC 94
Db	353 TCCGGAGAGTACACACACAAATGTGGGGGAGGGCGGCATTACGCTGAGTGGCGGCANAA 412
OY	95 CACCAACCTCACCGAGGAAGAGTGTGCGCCACCATCAGTACCTGATGGTGGCCCTGCACGA 154
Db	413 GCAGGCGTGGCCATTTGGGGCGCGCTGTGAAGAACCACCAAGCTCTTCTCTGGACGA 472
OY	155 GGGCCACACACAGATGACCGCCCGCGCGCGTGTGAGGTGGCGGTGATGTGACGACAT 214
Db	473 GGCACACGTGGCGCTGACAGCCGCCACGACGAGCGGTGTGTCAGCGCGCGCTGACCGCAT 532
RESULT 10	
LOCUS	BM485533
DEFINITION	682 bp mRNA linear EST 07-FEB-2002
ACCESSION	BM485533
VERSION	BM485533.1 GI:18606186

	KEYWORDS	EST.
	SOURCE	Chicken
	ORGANISM	Gallus gallus
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
		Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
		Phasianinae; Gallus.
	REFERENCE	1 (bases 1 to 682)
	AUTHORS	Cogburn,L.A. and Monsonego-Ornan,E.
	TITLE	ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library, USDA/IAPPS Animal Genome Project
JOURNAL COMMENT		Unpublished (2002)
	Contact:	Larry A. Cogburn
		University of Delaware
		Townsend Hall, Newark, DE 19717, USA
		Tel: 302-831-1335
		Fax: 302-831-2822
	Email:	cogburn@udel.edu, www.chickest.udel.edu.
FEATURES		Location/Qualifiers
source		1..682
	/organism=	"Gallus gallus"
	/strain=	"Commercial broiler and Ottawa Res. Centre Strains 90.6.21"
	/db_xref=	"taxon:9031"
	/clone=	"pgmln.pk001.o6"
	/clone_1lb=	"Normalized Chicken Breast Muscle, Leg Muscle, and Epiphyseal Growth Plate cDNA library (pgmln)"
	/sex=	"Male and Female"
	/tissue.type=	"Breast muscle, leg muscle and epiphyseal growth plate"
	/dev_stage=	"Breast,leg:Embryo(d19);post-hatch(1d,1.3,5.7,5.5,7.1,1 weeks);growth plate(1d,7d,14d post-hatch)"
	/lab_host=	"E. coli EMDH10B"
	/note=	"Vector: pCMVSPORT6; Library made from equivalent pools of total RNA isolated from each tissue (embryonic muscle 33.3%; juvenile muscle 33.3% and epiphyseal growing plate 33.3% of the final RN pool). Single pass sequencing from 5'-end"
BASE COUNT	151 a	179 c 230 g 120 t 2 others
ORIGIN		
Query Match	20.6%	Score 44; DB 13; Length 682;
Best Local Similarity	53.5%	Pred. No. 6.3'
Matches	92: Conservative	0; Mismatches 80; Indels 0; Gaps 0;
Oy	41 CAAGTGAACAAGAAGCTGGCGCTGGGGGCACCAATCCGCTCAGGTGACCACGCCAC	100
Db	162 CACGCTGATGAGAAAGTCGGGCCAAMACCCCACCAAAGAGCATGTCATGCATCATGCA	221
Oy	101 CCTCACCGAGAAAGCGTCGCGCCACCATCAGTAAGTAAGTGCGGCCCTGCACAGAGGCCA	160
Db	222 GGAGGTGACACGATGATGGACAGCGACCATCATCTTCACAGGATTCTCTGTGATATGTT	281
Oy	161 GACCAACATGACCCCGCGCGCGCGGTGAGAGTGGCGGTGATGTGAGACGAC	212
Db	282 CGCGCAGATGAAGAAGAGACGCCAACGAGCAAGTCTGAGAGAGAGCTGGCCAC	333
RESULT 11		
LOCUS	BFB64011	783 bp mRNA linear EST 19-JAN-2001
DEFINITION	963048D06.y1 C. reinhardtii CC-1690, stress condition I, normalized	
ACCESSION	BFB64011	, Lambda zap II Clamlydomonas reinhardtii cDNA, mRNA sequence.
VERSION	BFB64011.1	GI:12254155
KEYWORDS	EST.	
SOURCE	Clamlydomonas reinhardtii.	
ORGANISM	Clamlydomonas reinhardtii	
	Eukaryote; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;	
	Chlamydomonadaceae; Chlamydomonas.	
REFERENCE	1 (bases 1 to 783)	
AUTHORS	Grossman,A., Davies,J., Federspiel,N., Harris,E., Hauser,C., Lefebvre,P., Medermott,J.P., Shrago,J., Silflow,C. and Stern,D.	

TITLE	Analyses of the Chlamydomonas reinhardtii Genome: A Model,		
JOURNAL	Unicellular System for Analyzing Gene Function and Regulation in		
COMMENT	Vascular Plants; project phase 3		
CONTACT	Unpublished (2000)		
CONTACT	Contact: Charles Hauser		
DCMB	Box 91000		
UNIVERSITY	Duke University		
DURHAM	NC 27708-1000		
TEL	919 613 8159		
FAX	919 613 8177		
EMAIL	chauser@duke.edu		
FEATURES	Location/Qualifiers		
SOURCE	1. 783		
ORGANISM	"Chlamydomonas reinhardtii"		
STRAIN	"CC-1690 wild type mt+ 219t"		
DB_XREF	"taxon:3055"		
CLONE_LIB	"C. reinhardtii CC-1690, Stress condition 1, normalized, Lambda Zap II"		
NOTE	"vector: pBluescript II SK-; Site.1: EcoRI; Site.2: XhoI; This library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min, 1hr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, 1hr, 4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with Exassist (Stratagene) phase. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."		
BASE_COUNT	167 a 226 c 256 g 132 t 2 others		
ORIGIN			
Query Match	20.6%; Score 44; DB 12; Length 783;		
Best Local Similarity	52.8%; Pred. No. 6.4;		
Matches	95; Conservative 0; Mismatches 85; Indels 0; Gaps 0;		
QY	35 TCGGTACAAAGTGAACAAGAGCTGGGCGTGGCGGACCAATCCGGCTCAGTGACAC 94		
DB	386 TCCGGAGAAAGTACACACAAATGTGGCGAGGCGGCAATTCAGCTGAGTGGCGGCCAAA 445		
QY	95 CACCAACCCMCACCGGAGAAAGCTGTGCGCCACCATCGAGTACCTGTGCGCTCAGCA 154		
DB	446 GCAGCGCGGTGGCCATTGGCGCGCGGTGTGTAAGAACCCCAAGCTCTTCTCGAGCA 505		
QY	155 GGGCCAGACCAAGATGACCGCCCGCGCGCGCTCAGGCGCGGTGGATGTGGAGACAT 214		
DB	506 GGCACACGTGGCGGTAGACCGCCCGCAGCAGCGCGGTGTGTGACGCGCGCTGACCGAT 565		
RESULT 12			
LOCUS	B1995346 568 bp mRNA linear EST 25-OCT-2001		
DEFINITION	Lambda zap II Chlamydomonas reinhardtii CC-1690, Stress II (normalized),		
ACCESSION	B1995346		
VERSION	B1995346.1 GI:16430136		
KEYWORDS	EST.		
SOURCE	Chlamydomonas reinhardtii.		
ORGANISM	Chlamydomonas reinhardtii		
REFERENCE	Eukaryote: Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;		
AUTHORS	Chlamydomonadaceae; Chlamydomonas.		
TITLE	1 (bases 1 to 568)		
JOURNAL	Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre,		
COMMENT	P., McDermott,J.P., Shreager,J., Silflow,C. and Stern,D.		
DCMB	Analyses of the Chlamydomonas reinhardtii Genome: A Model,		
UNIVERSITY	Unicellular System for Analyzing Gene Function and Regulation in		
DURHAM	Vascular Plants. Project: 1031		
TEL	Unpublished (2001)		
FAX	Contact: Charles Hauser		
EMAIL	Box 91000		
CONTACT	Duke University		

Durham, NC 27708-1000
 Tel: 919 613 8159
 Fax: 919 613 8177
 Email: chauser@duke.edu.
 Location/Qualifiers

FEATURES

source

1.568

/organism="Chlamydomonas reinhardtii"
 /strain="CC-1690 wild type mt+ 21gr"
 /db_xref="taxon:3055"
 /clone_11b="C. reinhardtii CC-1690, Stress II (normalized)

/note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2: XhoI; Stress condition II library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (NH₄⁺ - containing) and shifted to TAP - NO₃ - (24hrs); H2 production conditions (0, 12hr, 24hr) see Mellis et al., (2000) Plant Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP + sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr). PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

BASE COUNT 80 a 206 c 185 g 97 t

ORIGIN

Query Match 20.5%; Score 43.8; DB 13; Length 568;
 Best Local Similarity 54.0%; Pred. No. 6.7;
 Matches 115; Conservative 0; Mismatches 92; Indels 6; Gaps 1;

OY 2 CAAGGAGAGCGCTGATCTGGCCCGCTGGTGCAGTGAACAAGAGCTGGG 61
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 251 CAAGGAGAGCGCTGATCTGGCCCGCTGGTGCAGTGAACAAGAGCTGGG 310
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 62 CTTGGGGCGGACCAATCCGCTCAAGTACACCAACCTTCACGAGAAGAGCTGT 121
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 311 CTTG-----CCGTACGAGGAGACCGCGGCTCTCCCGCTGCTCTACCACTG 364
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 122 CCGCCACATCGATGCTGTCGCGCTGCAGAGCGCCAGACGATGACCGCCCGG 181
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 365 GCG 424
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 182 CCGGCTCGAGGTGCGCGGTGATGTGACGACAT 214
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 425 CTGCGCGCAGACAGCCCGTGGCGCTTTCGGCAT 457
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13

LOCUS BG321153

DEFINITION Zm04_05f06_R Zm04_AAF-C_ECORC_cold_stressed_maize_seedlings Zea mays

ACCESSION BG321153 710 bp mRNA linear EST 27-FEB-2001

VERSION BG321153 Zm04_05f06, mRNA sequence.

KEYWORDS EST.

SOURCE Zea mays.

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 710)
 Singh, J.A., Wakul, K., Couroux, P., De Moors, A., Harris, L.J., Hattori, J.I., Ouellet, F., Robert, L.S., Spott, D., and Tinker, N.A.
 Expressed Sequence Tags from Cold-Stressed Maize Seedlings
 Unpublished (2001)
 Contact: Singh, J.A.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-Food Canada
 KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A

OC6, Canada
 Tel: (613) 759-1662
 Fax: (613) 759-1701
 Email: singh@em.agr.ca.
 Location/Qualifiers

FEATURES

source

1.710

/organism="Zea mays"
 /cultivar="C0328"
 /db_xref="taxon:4577"
 /clone="Zm04_05f06"
 /clone_11b="Zm04_AAF-C_ECORC_cold_stressed_maize_seedlings"

/note="Vector: Bluescript SK-/XhoI-EcoRI; Site_1: Eco RI; Site_2: Xho I; Lower temperature 50 C / hour from 22 to 120C; bring to 50 in 1 hour from 120C. Leave at 50C 2 days , photoperiod 16 hours. Light intensity was 125 uE-1. Library prepared by in vivo mass excision from amplified library."

BASE COUNT 118 a 263 c 213 g 80 t 36 others

ORIGIN

Query Match 20.3%; Score 43.4; DB 12; Length 710;
 Best Local Similarity 55.6%; Pred. No. 8.5;
 Matches 80; Conservative 1; Mismatches 63; Indels 0; Gaps 0;

OY 18 GATCTGGCCCGCTGGTTCGTCAGTGAACAAGAGCTGGCCCTGGCGGACCAAT 77
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 355 GAGAGGAGCNCATGCGCTGTCATGAGGCGCGCATCCAGAMCTGGCAGGGTCCC 414
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 78 CCGCTCAGGTGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 137
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 415 ATGAGAGAGATGAGAGCGCGCCCAAGCTTCAACCAAGGACGACGTGCTCAACTG 474
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 138 CTGCTGCGCTGTCAGAGAGGCCAG 161
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 475 CTGCGCTCTTTCACCGCGCGCAAG 498
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14

LOCUS AY109418

DEFINITION Zea mays CL2386_1 mRNA sequence.

ACCESSION AY109418 1856 bp mRNA linear HTC 25-MAY-2002

VERSION AY109418.1 GI:21213135

KEYWORDS HTC.

SOURCE Zea mays.

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 1856)
 Arthur, L.W., Hanafey, M., Morgan, M., and Tingey, S.V.
 Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes
 Unpublished (2002)
 2 (bases 1 to 1856)
 Coe, E.C.
 Direct Submission
 Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
 Location/Qualifiers

1.1856

/organism="Zea mays"
 /db_xref="MaizeDB:630950"
 /db_xref="taxon:4577"
 /clone="CL2386_1"
 /clone_11b="Maize Mapping Project/Dupont Consensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the

Query Match	20.28;	Score 43.2;	DB 11;	Length 1856;
Best Local Similarity	56.28;	Pred. No. 11;		
Matches 81; Conservative	0;	Mismatches 63;	Indels 0;	Gaps 0;

RESULT 15	LOCUS	DEFINITION	ACCSSION	BE754521	558 bp	linear	EST 25-APR-2001
	LOCUS	BE754521	208060	MARC 2B0V Bos	taurus	CDNA 5', mRNA sequence.	

LOCUS	558 bp	linear	EST 25-APR-2001
DEFINITION	208060 MARC 2BOV Bos taurus	5', mRNA	
ACCESSION	BE754521	CDNA	sequence.
VERSION	BE754521.1	GI:10168513	
KEYWORDS	EST.		
SOURCE	COW.		
ORGANISM	Bos taurus		

REFERENCE
AUTHORS
1 (bases 1 to 558)
Smith, T. P. L., Grosse, W. M., Freking, B. A., Roberts, A. J., Stone, R. T
Lacini, P. A., Metacora, Choriata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.

TITLE	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL	Genome Res. 11 (4), 626-630 (2001)
MEDLINE	21180013
COMMENT	Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel.: 402 762 4366
Tel.: 402 762 4360
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -mismore 18
and -mismatch 12 options.
PCR primers
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCGAGTCACGACG
Plate: 54 Row: M Column: 18
Seq primer: ATTAGTGACACCTATAG.

BASE COUNT	FEATURES	SOURCE
113 a	Location/Qualifiers	1. .558
166 c	/organism="Bos taurus"	
	/db_xref="taxon:9913"	
	/clone_11b="MARC 2B0V"	
	/tissue_type="pooled"	
	/lab_host="DH10B"	
	/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI; library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."	
172 g		107 t

Query Match 20.1%; Score 43; DB 12; Length 558;

	Best Local Similarity	54.0%	Pred. No. 10:	
Matches	88	Conservative	0	Mismatches 75; Indels 0; Gaps 0:
QY	25	CCCCGGTGGGTGCTACAGGTGAACAAGATGGCGCTTGGCGGACCAATCGGCTC	84	
Db	198	CCATCATGGGCCAGTTCTGAGCACCCCAATATATCTCCGCTGGAGGGGCTGTCCACCAACA	257	
QY	85	AGGTGACCAACCAACCAACCGAGGAAGAGCTGTGCGACACATGAGTACTCTGTGC	144	
Db	258	GCCTGCCCTCTATGATCTCTACCCGAGTTCATGAGGAACGGTGCCTGTGACTCTTCTCTC	317	
QY	145	GCCTGCACGAGGCGACAGACAGATAGCGGCCCGCGCGCGCT	187	
Db	318	GGCTGAACGAGGGCCAGTTCAACCGTGAATCCAGCTCTGTGGCGAT	360	

Search completed: November 13, 2002, 04:00:40
Job time : 1154.63 secs

Search completed: November 13, 2002, 04:00:40
Job time : 1154.63 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 ; Search time 21.7711 Seconds

(without alignments)
3487.380 Million cell updates/sec

Title: US-09-697-123b-21

Perfect score: 214

Sequence: 1 tcaagagagagcgcacgat.....ccggtgatgtgacacacat 214

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications -NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCIT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCITUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80.4	37.6	5096	10 US-09-984-711-5	Sequence 5, Appl1
2	80.4	37.6	5099	9 US-10-075-460-5	Sequence 1, Appl1
3	80.4	37.6	5099	10 US-09-887-052-1	Sequence 3, Appl1
4	80.4	37.6	5099	10 US-09-887-052-3	Sequence 3, Appl1
5	80.4	37.6	5099	10 US-09-887-052-5	Sequence 3, Appl1
6	41	19.2	399	10 US-09-878-574-3598	Sequence 3598, Ap
7	38.6	18.0	399	10 US-09-861-893-29	Sequence 29, Appl
8	38.6	18.0	431	10 US-09-954-456-1034	Sequence 1034, Ap
9	38.2	17.9	1818	10 US-09-815-242-7989	Sequence 7989, Ap
10	37.2	17.4	250	10 US-09-764-853-229	Sequence 229, App
11	37.2	17.4	2383	10 US-09-925-302-317	Sequence 317, App
12	37	17.3	879	10 US-09-861-289-11	Sequence 11, Appl
13	37	17.3	13613	10 US-09-861-289-3	Sequence 3, Appl1
14	36.8	17.2	855	10 US-09-815-242-7681	Sequence 7681, Ap
15	36.6	17.1	350	10 US-09-815-242-7681	Sequence 3310, A
16	36	16.8	6252	10 US-09-864-824-313	Sequence 313, App
17	35.6	16.6	429	10 US-09-815-242-4133	Sequence 4133, Ap
18	35.6	16.6	1680	10 US-09-820-721A-2	Sequence 2, Appl1
19	35.4	16.5	409	10 US-09-878-574-3856	Sequence 3856, Ap

20	35.4	16.5	3057	10 US-09-815-242-4131	Sequence 4131, Ap
21	35.2	16.4	2712	10 US-09-748-033-4	Sequence 4, Appl1
22	35.2	16.4	3435	10 US-09-803-670-1	Sequence 1, Appl1
23	35.2	16.4	8868	10 US-09-803-670-3	Sequence 3, Appl1
24	35	16.4	3945	10 US-09-921-771-4	Sequence 4, Appl1
25	35	16.4	3969	10 US-09-982-610-23	Sequence 23, Appl
26	35	16.4	7419	10 US-09-815-242-4009	Sequence 4009, Ap
27	34.8	16.3	403	10 US-09-878-574-105	Sequence 2105, Ap
28	34.8	16.3	1101	10 US-09-815-242-7943	Sequence 7943, Ap
29	34.8	16.3	1416	10 US-09-815-242-7943	Sequence 4160, Ap
30	34.8	16.3	1476	10 US-09-997-664-99	Sequence 99, Appl
31	34.8	16.3	6491	10 US-09-732-680A-1	Sequence 1, Appl1
32	34.6	16.2	1539	10 US-09-732-680A-1	Sequence 105, App
33	34.6	16.2	1914	10 US-09-815-242-7960	Sequence 777, App
34	34.6	16.2	2090	10 US-09-822-887-3	Sequence 1, Appl1
35	34.6	16.2	2142	10 US-09-822-887-3	Sequence 105, App
36	34.6	16.2	2493	10 US-09-822-887-1	Sequence 1, Appl1
37	34.6	16.2	4826	10 US-09-772-304A-1	Sequence 9, Appl1
38	34.4	16.1	2329	10 US-09-816-828-9	Sequence 367, App
39	34.2	16.0	1467	10 US-09-887-576-777	Sequence 5, Appl1
40	34	15.9	480	12 US-10-052-586-367	Sequence 10, Appl1
41	33.8	15.8	11575	10 US-09-938-956-5	Sequence 7949, Ap
42	33.8	15.8	12341	12 US-10-033-190-5	
43	33.8	15.8	13737	9 US-10-074-279-10	
44	33.8	15.8	14446	9 US-09-810-861B-4	
45	33.6	15.7	801	10 US-09-815-242-7949	

ALIGNMENTS

RESULT 1

US-09-984-711-5
Sequence 5, Application US/09984711

Patent No. US20020119549A1

GENERAL INFORMATION:

APPLICANT: MOECKEL, Bettina

APPLICANT: BATHÉ, Brigitte

APPLICANT: STEPHAN, Hans

APPLICANT: KREUTZER, Caroline

APPLICANT: HERMANN, Thomas

APPLICANT: PFEFFERLE, Walter

APPLICANT: BINDER, Michael

TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE TP51 GENE

FILE REFERENCE: 204209US0

CURRENT APPLICATION NUMBER: US/09/984, 711

CURRENT FILING DATE: 2001-10-31

PRIOR APPLICATION NUMBER: DE10108230.9

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn version 3.1

SEQ ID NO 5

LENGTH: 5096

TYPE: DNA

ORGANISM: Corynebacterium glutamicum

FEATURE:

NAME/KEY: CDS

LOCATION: (702)..(4196)

OTHER INFORMATION:

US-09-984-711-5

Query Match 37.6%: Score 80.4; DB 10; Length 5096;
Best Local Similarity 66.08; Pred. No. 1.3e-11;

Matches 136; Conservative 0; Mismatches 61; Indels 9; Gaps 1;

QY 9 AAGCGTACGATCTGGCCGCGTGGTGTGATACAGGTGAAGAAGCTGGCGCTGGGC 68
DB 1578 AAGCGTACGATCTGGCCGCGTGGTGTGATACAGGTGAAGAAGCTGGCGCTGGGC 1637
QY 69 GGCACCAATCGCGGTGAGTACACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 128
DB 1638 GGCACCAATCGCGGTGAGTACACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1688

OY 129 ATCGAGTACCTGGTGGCCCTGCACGAGGCCAGACCAGCATGACCGCCCGCGCGCTC 188
|||||
Db 1689 ATCGAGTACCTGGTGGCTGTGCACGAGGTGACCGCGTGTACTTCTCCAAATGGTGAA 1748
OY 189 GAGGTCCCGGTGATGTGACGACAT 214
|||
Db 1749 GAGATCCCACTGACGACCATGACAT 1774

RESULT 2

US-10-075-460-5
; Sequence 5, Application US/10075460
; Patent No. US2002015557A1
; GENERAL INFORMATION:
; APPLICANT: MOCKEL, BETTINA
; APPLICANT: BATHE, BRIGITTE
; APPLICANT: HANS, STEFAN
; APPLICANT: KREUTZER, CAROLINE
; APPLICANT: HERMANN, THOMAS
; APPLICANT: PFEFFERLE, WALTER
; APPLICANT: BINDER, MICHAEL
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE
; FILE REFERENCE: 218472US0X
; CURRENT APPLICATION NUMBER: US/10/075.460
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: DE 10107230.9
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: DE 10162386.0
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
; OTHER INFORMATION:
US-10-075-460-5

Query Match 37.6%; Score 80.4; DB 9; Length 5099;
Best Local Similarity 66.0%; Pred. No. 1.3e-11;

Matches 136; Conservative 0; Mismatches 61; Indels 9; Gaps 1;

OY 9 AAGCGTACGATCTGGCCCGCGGTGGTGTACAGGTGACGAAGAAGCTGGCGCTGGGC 68
|||||
Db 1578 AAGCGTACGACCTGGCTCGCGTGTGCTTACAGATCAACCGCAAGCTCGCGCTTGGT 1637
OY 69 GGCACCAATCCGGCTCAGGTGACCCACACCCCTCAGGAGGAAGAGCTGCGCCACC 128
|||||
Db 1638 GGCACCAACGATGTTGTGATGACT-----CTTACTGAAGAGAGCATCCCAACACC 1688
OY 129 ATCGAGTACCTGGTGGCCCTGCACGAGGCCAGACCATGACCGCCCGCGCGCGCTC 188
|||||
Db 1689 ATCGAGTACCTGGTGGCTGTGCACGAGGTGACGCGGTGATGACTTCTCCAAATGGTGAA 1748
OY 189 GAGGTCCCGGTGATGTGACGACAT 214
|||
Db 1749 GAGATCCCACTGACGACCATGACAT 1774

RESULT 3

US-09-887-052-1
; Sequence 1, Application US/09887052
; Patent No. US20020119537A1
; GENERAL INFORMATION:
; APPLICANT: MOCKEL, Bettina
; APPLICANT: BATHE, Brigitte
; APPLICANT: HERMANN, Thomas
; APPLICANT: PFEFFERLE, Walter
; APPLICANT: BINDER, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE

; FILE REFERENCE: 204212US0X
; CURRENT APPLICATION NUMBER: US/09/887.052
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: DE10107229.5
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
US-09-887-052-1

Query Match 37.6%; Score 80.4; DB 10; Length 5099;
Best Local Similarity 66.0%; Pred. No. 1.3e-11;
Matches 136; Conservative 0; Mismatches 61; Indels 9; Gaps 1;

OY 9 AAGCGTACGATCTGGCCCGCGGTGGTGTACAGGTGACGAAGAAGCTGGCGCTGGGC 68
|||||
Db 1578 AAGCGTACGACCTGGCTCGCGTGTGCTTACAGATCAACCGCAAGCTCGCGCTTGGT 1637
OY 69 GGCACCAATCCGGCTCAGGTGACCCACACCCCTCAGGAGGAAGAGCTGCGCCACC 128
|||||
Db 1638 GGCACCAACGATGTTGTGATGACT-----CTTACTGAAGAGAGCATCCCAACACC 1688
OY 129 ATCGAGTACCTGGTGGCCCTGCACGAGGCCAGACCATGACCGCCCGCGCGCGCTC 188
|||||
Db 1689 ATCGAGTACCTGGTGGCTGTGCACGAGGTGACGCGGTGATGACTTCTCCAAATGGTGAA 1748
OY 189 GAGGTCCCGGTGATGTGACGACAT 214
|||
Db 1749 GAGATCCCACTGACGACCATGACAT 1774

RESULT 4

US-09-887-052-3
; Sequence 3, Application US/09887052
; Patent No. US20020119537A1
; GENERAL INFORMATION:
; APPLICANT: MOCKEL, Bettina
; APPLICANT: BATHE, Brigitte
; APPLICANT: HERMANN, Thomas
; APPLICANT: PFEFFERLE, Walter
; APPLICANT: BINDER, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE
; FILE REFERENCE: 204212US0X
; CURRENT APPLICATION NUMBER: US/09/887.052
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: DE10107229.5
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
US-09-887-052-3

Query Match 37.6%; Score 80.4; DB 10; Length 5099;
Best Local Similarity 66.0%; Pred. No. 1.3e-11;
Matches 136; Conservative 0; Mismatches 61; Indels 9; Gaps 1;

OY 9 AAGCGTACGATCTGGCCCGCGGTGGTGTACAGGTGACGAAGAAGCTGGCGCTGGGC 68
|||||
Db 1578 AAGCGTACGACCTGGCTCGCGTGTGCTTACAGATCAACCGCAAGCTCGCGCTTGGT 1637
OY 69 GGCACCAATCCGGCTCAGGTGACCCACACCCCTCAGGAGGAAGAGCTGCGCCACC 128


```
Db 1638 GCGGACACGATGTTGATGACT-----CTTACGAAGAGACATCGCAACACC 1688
QY 129 ATGAGTACCTGGTGGCTGACGAGGCGCCAGCAGATGACCGCCCGCGGCGTC 188
Db 1689 ATGAGTACCTGGTGGCTGACGAGGCGCGCTCATGACTTCTCAAAATGGTGA 1748
QY 189 GAGTGGCGGTGATGAGCAGCAT 214
Db 1749 GAGATCCCAAGTCAGACGATGACAT 1774
```

RESULT 5

```
US-09-887-052-5
; Sequence 5, Application US/09887052
; Patent No. US20020119537A1
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: BATHÉ, Brigitte
; APPLICANT: HERMANN, Thomas
; APPLICANT: PEPPERLE, Walter
; APPLICANT: BINDER, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE PROB GENE
; FILE REFERENCE: 204212US0X
; CURRENT APPLICATION NUMBER: US/09/887,052
; PRIOR FILING DATE: 2001-08-25
; PRIOR APPLICATION NUMBER: DE10107229.5
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
US-09-887-052-5
```

```
Query Match 37.6%; Score 80.4; DB 10; Length 5099;
Best Local Similarity 66.0%; Pred. No. 1.3e-11;
Matches 136; Conservative 0; Mismatches 61; Indels 9; Gaps 1;
```

```
QY 9 AAGCCCTACGATCTGGCCGCGTGGTGGTACAGAGTGAACAAGAGTGGCGCTGGGC 68
Db 1578 AAGCCCTACGATCTGGCTGGCTGGTGGTGGTACAGATCAAGCAAGCTGGCGCTGGT 1637
QY 69 GGCACCAATCCGCGCTGAGGTGACCAACCACCACTTCACGAGGAAGAGTGGCGGCGAC 128
Db 1638 GGCACCAATCCGCGCTGAGGTGACCAACCACCACTTCACGAGGAAGAGTGGCGGCGAC 1688
QY 129 ATCGAGTACCTGGTGGCTGGCTGGCAAGAGGCGCCAGACGATGACCGCCCGCGGCGTC 188
Db 1689 ATCGAGTACCTGGTGGCTGGCTGGCAAGAGGCGCGTCAATGCTCAAAATGGTGA 1748
QY 189 GAGTGGCGGTGATGAGCAGCAT 214
Db 1749 GAGATCCCAAGTCAGACGATGACAT 1774
```

RESULT 6

```
US-09-878-574-3598
; Sequence 3598, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Bytun, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
```

```
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 3598
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-008-Q1-B1-C6
US-09-878-574-3598
```

```
Query Match 19.2%; Score 41; DB 10; Length 392;
Best Local Similarity 56.2%; Pred. No. 0.051;
Matches 77; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
```

```
QY 76 ATCCGCTCAGGTGACCAACCACTGACGAGGAAGAGTGGTGGCCACCATGAGT 135
Db 2 ACCTGCGCCAGGCTTCCGCTCATGACCGCCGACACGAGGCGTGTACGCGCGCAGG 61
QY 136 ACCTGCTGCGCTGACAGAGGCGCCAGACCATGACCGCCCGCGGCGTGGAGTGC 195
Db 62 ACCTGAGGCGCTCTCTCAGTGGCTCGGCGCGTGGCGCGTGGCGCCGACGAGTGGGCTCA 121
QY 196 CGGTGATGTGAGCGAC 212
Db 122 TGCTGCGGAGGTGAC 138
```

RESULT 7

```
US-09-861-893-29
; Sequence 29, Application US/09861893
; Patent No. US20020045257A1
; GENERAL INFORMATION:
; APPLICANT: Feinberg, Andrew
; APPLICANT: Strichman-Almashanu, Liora
; APPLICANT: Jiang, Shan
; TITLE OF INVENTION: METHODS FOR ASSAYING GENE IMPRINTING AND
; FILE REFERENCE: 01107,00128
; CURRENT APPLICATION NUMBER: US/09/861,893
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 60/206,158
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/206,161
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-861-893-29
```

```
Query Match 18.0%; Score 38.6; DB 10; Length 399;
Best Local Similarity 52.1%; Pred. No. 0.2;
Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
```

```
QY 36 CGGTCAAGGTGAACAAGAGTGGGCGGCGGACCAATCCGGCTAGTGAACACC 95
Db 224 CAGACCAAGTGTCTCAACATGCGCAACGAGACTGATCAAGGTGGCGCTGTCAAC 283
QY 96 ACCACCTTACCGAGAGAGTGTGCGCACCATGAGTACTGCTGCGCTGACGAG 155
Db 284 AACGCTTCTACTGTGAGACCTGCACTTCACCATGAGAGGAGACGACTACTTC 343
QY 156 GGCAGACACGATGAGACCGCCCGCGGCGCTGAGAGTCCCGGTG 200
Db 344 ATCAAGACACACGCGCCGAGAGACCTGGGACGCTGCGCTG 388
```

RESULT 8

```
US-09-954-456-1034/C
; Sequence 1034, Application US/09954456
; Patent No. US20020115057A1
```

```

: GENERAL INFORMATION:
: APPLICANT: Young, Paul
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
: TITLE OF INVENTION: Sets
: FILE REFERENCE: 689290-76
: CURRENT APPLICATION NUMBER: US/09/954,456
: CURRENT FILING DATE: 2001-09-18
: PRIOR APPLICATION NUMBER: US/60/233,617
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234,052
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,923
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,134
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,637
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,638
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,711
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,720
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,840
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,863
: NUMBER OF SEQ ID NOS: 2276
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1034
: LENGTH: 431
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-954-456-1034

Query Match          18.0%: Score 38.6: DB 10: Length 431:
Best Local Similarity 52.1%: Pred. No. 0.2:
Matches 86: Conservative 0: Mismatches 79: Indels 0: Gaps 0:

QY 36 CGGTACAGAGTGAACAAGAGCTGGGCGTGGCGGACCAATCCGCTCAGGTGACGACC 95
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 177 CAGACCAACGTGCTGCATCAATCCCAAGAGAGACTGCATCAAGTGGCGCGCTGCTAC 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 96 ACCACCTCAGCGAGAGAGCGTGGTGGCCACCATCGAGTACGTGTCGCCCTGCACAG 155
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 117 AACGCTTTTACCTGAGAGAACTGCATCTACCATCGAGGCGCAAGACACGCACTACTTC 58
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 156 GGGCAGACGACGATGACCGCCCGCGGCGGTGAGGTGCGCGT 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 57 ATCAAGACCACGACCGCCGAGAGCGACGCGGCGGCGGCGGTTG 13

RESULT 9
US-09-815-242-7989
: Sequence 7989, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: TITLE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
```

```

: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7989
: LENGTH: 1818
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(1818)
US-09-815-242-7989

Query Match          17.9%: Score 38.2: DB 10: Length 1818:
Best Local Similarity 53.3%: Pred. No. 0.29:
Matches 104: Conservative 0: Mismatches 88: Indels 3: Gaps 1:
```

```

QY 23 GGGCCGCGTGGGTGCGGTACAGGTGAACAAGAGCTGGGCGGCGGACCAATCCGGC 82
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 354 GGGCGGCGTGGCGTCCGATGCTGTGAGTAAGATCGACCGCGGCGGCGGCGGCGGA 413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 83 TCAGTGACCAACCAACCTCAGCAG--GAGAGCTGCTGGCCACCATGAGTACT 139
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 414 CTGGGTATCGACGACGATCTTGACCTGTTCGACAACTCGGCGGCGACGAGCACT 473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 140 GGTGCCCTGTCAGAGAGGCGACACGATGACCGCGCGCGCGCGCGTCCAGTCCGGT 199
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 474 CGACTTCGCGATGCTGTACGCGAGCGCCCTGACGCGCATCCGCGCTCGACGAGAA 533
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 200 GATGTGACGACAT 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 534 GATGACGACACAT 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

RESULT 10
US-09-764-853-229/c
: Sequence 229, Application US/09764853
: Patent No. US20020090672A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: P1206
: CURRENT APPLICATION NUMBER: US/09/764,853
: PRIOR APPLICATION NUMBER: 2001-01-17
: PRIOR APPLICATION data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 939
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 229
: LENGTH: 250
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-764-853-229
```

```

Query Match          17.4%: Score 37.2: DB 10: Length 250:
Best Local Similarity 56.9%: Pred. No. 0.42:
Matches 66: Conservative 1: Mismatches 49: Indels 0: Gaps 0:

QY 13 GCTACGATCTGCGCCGCGTGGCTGCTACAGGTGAACAAGAGCTGGCGGCGCA 72
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 125 GCTCACAAGTAGAGGCTGGGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 73 CCAATCCGGCTCAGGTGACCAACCAACCCCTCAGGAGGAAGAGCTGTGCGCAC 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 65 CCACTTTCCGAGAGCTCCACTCGACACCCMCCTGGAAGGCGTCTCCGCGCC 10
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 11
US-09-925-302-317
: Sequence 317, Application US/09925302
: Patent No. US2002004941A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA104
: CURRENT APPLICATION NUMBER: US/09/925.302
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05918
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 896
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 317
: LENGTH: 2383
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-925-302-317

Query Match
Best Local Similarity 17.4%; Score 37.2; DB 10; Length 2383;
Matches 96; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

OY 8 GAAGCGCTACGATCTGGCCCGCGTGGGTGCGTCAAGTGACAGAGAGCTGGGCTGGG 67
DB 838 GCACCACGAGAGCGCGGTCTGTAACCTCAGCAAGTCTCTGAGATGCTGGAGAGTG 897
OY 68 CGGCACCAATCCGGCTCAGGTGACCAACCACTCAGCAAGAGAGCTGCTGCCAC 127
DB 898 CGAGCGCGGATGCGAGCGCCGCTCAGCAACCCCTCAGCAAGAGTGTCCCGCAT 957
OY 128 CATGAGTACCTGTGGCGGCTGCGACGAGGCGCAGACATACCGCCCGCGCGCT 187
DB 958 CTTGGAACCTCGAGAGCAAGTACTATGAGAGTACAGATGTCCGACCTGTGGACCT 1017
OY 188 CGAGTCCCGGTG 201
DB 1018 TGTCTGTGGCCATCG 1031

RESULT 12
US-09-861-289-11
: Sequence 11, Application US/09861289
: Patent No. US20020110897A1
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OF INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600.438US1
: CURRENT APPLICATION NUMBER: US/09/861,289
: CURRENT FILING DATE: 2001-05-18
: PRIOR APPLICATION NUMBER: 09/105,537
: PRIOR FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 11
: LENGTH: 879
: TYPE: DNA
: ORGANISM: Streptomyces venezuelae
US-09-861-289-11

Query Match
Best Local Similarity 17.3%; Score 37; DB 10; Length 879;
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

OY 87 GTGACCAACCACTCAGCGAGAGAGTGTGCGCACCATCGATGCTGGCGC 146
```

```
DB 508 GTACCGGCTCTACTCTACGACAGAGCTGCTGACATCGCCAGAACATCGGCCC 567
OY 147 CTGACAGAGGCGGACAGACATGACCGCCCGCGCGCTGAGGTGCGGTGATGTG 206
DB 568 TCGCCCGCGCGGAGCTGAGATCAACGAGTCAACCGCTTACTCTGAGCGGCGCG 627
OY 207 GACGA 211
DB 628 GCCGA 632

RESULT 13
US-09-861-289-3/C
: Sequence 3, Application US/09861289
: Patent No. US20020110897A1
: GENERAL INFORMATION:
: APPLICANT: Sherman, D.H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OF INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600.438US1
: CURRENT APPLICATION NUMBER: US/09/861,289
: CURRENT FILING DATE: 2001-05-18
: PRIOR APPLICATION NUMBER: 09/105,537
: PRIOR FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3
: LENGTH: 13613
: TYPE: DNA
: ORGANISM: Streptomyces venezuelae
US-09-861-289-3

Query Match
Best Local Similarity 17.3%; Score 37; DB 10; Length 13613;
Matches 70; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

OY 87 GTGACCAACCACTCAGCGAGAGAGTGTGCGCACCATCGATGCTGGCGC 146
DB 9482 GTACCGGCTCTACTCTACGACAGAGCTGCTGACATCGCCAGAACATCGGCCC 9423
OY 147 CTGACAGAGGCGGACAGACATGACCGCCCGCGCGCTGAGGTGCGGTGATGTG 206
DB 9422 TCGCCCGCGCGGAGCTGAGATCACCAGATCACCAGTCAACCGCTTACTCTGAGCGGCGCG 9363
OY 207 GACGA 211
DB 9362 GCCGA 9358

RESULT 14
US-09-815-242-7681/C
: Sequence 7681, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlson, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
```

```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,635
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7681
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (855)
US-09-815-242-7681
```

```
Query Match          17.2%  Score 36.8; DB 10; Length 855;
Best Local Similarity 50.0%; Pred. No. 0.59;
Matches 92; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
```

```

OY 28 GCGTGGTCCGTACAGGTGACACAGAGCTGGGCTGGGCGGACCAATCCGCTCAGG 87
DB 500 GCGTGGGCGCGCGCGCGGCTGACGTGACGACAGACGACGCGCTCGACGCGCTCGGCGAGC 441
OY 88 TGACACACACACACCTTCACCGAGAGAGAGCTGTGCGCCACCATCGAGTACTGTGGCCGC 147
DB 440 AGGCGATGGGCGACCTCGCGCGCGCGCGCGCGCGCGCGCATCTGACACAGCGGCTGGCGC 381
OY 148 TGCACGAGGCGCCAGACACAGATGACCGCCCGCGCGCGCTGAGTGGTGATGTGG 207
DB 380 GCGCGCTGGCGGCGCGCGCGCGGAAAGCCCTCGCGGAAAGCCAGCGCATGCTGTGTGG 321
OY 208 ACCGA 211
DB 320 CCGCA 317
```

```

RESULT 15
US-09-864-761-33210/c
; Sequence 33210, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33210
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005973.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.88
; OTHER INFORMATION: EST HUMAN HIT: AA45418.1, EVALU 2.20e-01
; OTHER INFORMATION: NT HIT: AF10599.1, EVALU 1.00e-127
; OTHER INFORMATION: SWISSPROT HIT: Q04844, EVALU 3.00e-22
US-09-864-761-33210
```

```
Query Match          17.1%  Score 36.6; DB 10; Length 350;
Best Local Similarity 50.9%; Pred. No. 0.61;
Matches 87; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
```

```

OY 24 GCGCGGTGGTGTGACAGGTGACACAGAGCTGGGCTGGGCGGACCAATCCGCT 83
DB 231 GCCCAGAGAGCGCGGCGACGCTCCAGAGAACCTGGGGCGAGGGGCGGCTTAGGGAC 172
OY 84 CAGGTGACCAACCAACCTCCACGAGAGAGAGCTGTGCGCCAGCATCGAGTACTGTG 143
DB 171 GAGGTAGTACAGAACGCCACCCGACCCGGGCTGACACGCGCCCTCGCGCTTAGCTGG 112
OY 144 GCGCTGACAGAGGCGCAACACCATGACCGCCCGCGGCGCTGAGCTG 194
DB 111 CCGACCGCGGGGACATGCTGTGGTGGGCGCTCCGCTGGACACGTTG 61
```

Search completed: November 12, 2002, 16:58:57
Job time : 36.7711 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:56:36 ; Search time 28.6983 Seconds
(without alignments)
2286.857 Million cell updates/sec

Title: US-09-697-123B-21
Perfect score: 214
Sequence: 1 tcaaggagagcgtacgacat.....ccggtgagtgtagcagacat 214

Scoring table:
IDENTITY_NIC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, NA: *
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/lna/PCrus.COMB.seq: *
6: /cgn2_6/ptodata/1/lna/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129.6	60.6	970	1 US-08-250-030-1	Sequence 1, Appl1
2	129.6	60.6	970	5 PCT-US95-06790-1	Sequence 1, Appl1
3	108.8	50.8	3447	2 US-08-313-185-57	Sequence 57, Appl1
4	108.8	50.8	3447	3 US-09-082-614A-57	Sequence 57, Appl1
5	42.2	19.7	1734	6 5352575-8	Patent No. 5352575
6	38.4	17.9	68750	3 US-09-335-409-1	Sequence 1, Appl1
7	38.4	17.9	68750	4 US-09-568-102-1	Sequence 1, Appl1
8	38.4	17.9	68750	4 US-09-567-969-1	Sequence 1, Appl1
9	38.4	17.9	68750	4 US-09-568-480-1	Sequence 1, Appl1
10	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
11	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
12	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
13	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
14	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
15	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
16	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
17	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
18	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
19	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
20	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
21	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
22	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
23	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
24	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
25	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
26	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1
27	38.4	17.9	68750	4 US-09-568-486-1	Sequence 1, Appl1

28	35.8	16.7	1176	4 US-09-479-409-34	Sequence 34, Appl1
29	35.8	16.7	1176	4 US-09-479-453-34	Sequence 34, Appl1
30	35.8	16.7	1176	4 US-08-911-853-29	Sequence 29, Appl1
31	35.8	16.7	1176	4 US-09-479-409-29	Sequence 29, Appl1
32	35.8	16.7	1176	4 US-09-479-453-29	Sequence 29, Appl1
33	35.6	16.6	1602	4 US-07-820-011A-1	Sequence 1, Appl1
34	35.6	16.6	1602	5 PCT-US93-00445-1	Sequence 1, Appl1
35	35.6	16.6	4403765	4 US-09-103-840A-2	Sequence 2, Appl1
36	35.6	16.6	4411529	4 US-09-103-840A-1	Sequence 1, Appl1
37	35.4	16.5	9515	1 US-08-920-812-13	Sequence 13, Appl1
38	35.4	16.5	9515	1 US-08-920-812-13	Sequence 13, Appl1
39	35.4	16.5	9515	1 US-08-921-177-13	Sequence 13, Appl1
40	35.4	16.5	9515	1 US-08-362-577C-13	Sequence 13, Appl1
41	35.4	16.5	9515	2 US-08-920-828-13	Sequence 13, Appl1
42	35.4	16.5	30001	1 US-08-125-468-1	Sequence 13, Appl1
43	35.4	16.5	30001	2 US-08-474-933-1	Sequence 1, Appl1
44	35.2	16.4	2712	3 US-09-025-691-4	Sequence 1, Appl1
45	35.2	16.4	71989	4 US-09-443-501A-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-08-250-030-1
Sequence 1, Application US/08250030
Patent No. 5643723
GENERAL INFORMATION:
APPLICANT: Persing, David H.
TITLE OF INVENTION: Detection of a Genetic Locus Encoding
TITLE OF INVENTION: Resistance to Rifampin in Mycobacterial Cultures and in
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg & Woessner
STREET: 3500 IDS Center
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/250,030
FILING DATE: 26-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Muelling, Ann M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 150,105US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-339-0331
TELEFAX: 612-339-3061
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 970 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-250-030-1
Query Match: 60.6%; Score 129.6; DB 1; Length 970;
Best local Similarity: 79.0%; Pred. No. 7,2e-21;
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;
OY 1 TCAAGGAGAGCCCTACGATCTGCGCCGCTGCTGCTACAGAGTGAACAAGAGCTCG 60
DB 26 TCAAGGAGAGCCCTACGATCTGCGCCGCTGCTGCTACAGAGTGAACAAGAGCTCG 85

QY	61	GCTGGGGGGCACCAATCGGGCTAGGAGACACACACACCTCACCGAGGAACGCTCG	120
Db	86	GCGTCGATGTCGGGCAAGCC-----CATCAGTCGTGTGACGCTGACCCGAAGAAGACTCG	139
QY	121	TGCGACCAATCGAGTACCTGGTGGCGCTGGACGAGGAGGCCAGACCAACGATGAGACGGCCCCG	180
Db	140	TGGCCACCAATCCAAATATTTGGTCCGTTGGACGAGGAGGTGACGACACGATGACCTTCGCG	199
QY	181	GCGGCGTTCGAGGTGCCGGTGGATGTGGGACACAT	214
Db	200	GCGGCGTTCGAGGTGCCGGTGGAAACGACGACAT	233

```

1      RESULT 2
2      PCT-US95-06790-1
3      : Sequence 1, Application PC/TUS9506790
4      : GENERAL INFORMATION:
5      :   APPLICANT: Mayo Foundation for Medical Education and Research
6      :   APPLICANT: and Hoffmann-La Roche Inc.
7      :   TITLE OF INVENTION: Detection of a Genetic Locus Encoding
8      :   TITLE OF INVENTION: Resistance to Ritampin
9      :   NUMBER OF SEQUENCES: 15
10     : CORRESPONDENCE ADDRESS:
11     :   ADDRESSEE: Schweegman, Lundberg & Woessner
12     :   STREET: 3500 IDS Center
13     :   CITY: Minneapolis
14     :   STATE: MN
15     :   COUNTRY: USA
16     :   ZIP: 55402
17     : COMPUTER READABLE FORM:
18     :   MEDIUM TYPE: Floppy disk
19     :   COMPUTER: IBM PC compatible
20     :   OPERATING SYSTEM: PC-DOS/MS-DOS
21     :   SOFTWARE: PatentIn Release #1.0, Version #1.25
22     :   CURRENT APPLICATION DATA:
23     :     APPLICATION NUMBER: PCT/US95/06790
24     :     FILING DATE: 26-MAY-1995
25     :     CLASSIFICATION:
26     :   ATTORNEY/AGENT INFORMATION:
27     :     NAME: Raasch, Kevin W.
28     :     REGISTRATION NUMBER: 35,651
29     :     REFERENCE/DOCKET NUMBER: 150.105WO1
30     :     TELECOMMUNICATION INFORMATION:
31     :       TELEPHONE: 612-339-0331
32     :       TELEFAX: 612-339-3061
33     :   INFORMATION FOR SEQ ID NO: 1:
34     :     SEQUENCE CHARACTERISTICS:
35     :       LENGTH: 970 base pairs
36     :       TYPE: nucleic acid
37     :       STRANDEDNESS: single
38     :       TOPOLOGY: linear
39     :     MOLECULE TYPE: DNA
40     :   :
41     :   :
42     :   :
43     :   :
44     :   :
45     :   :
46     :   :
47     :   :
48     :   :
49     :   :
50     :   :
51     :   :
52     :   :
53     :   :
54     :   :
55     :   :
56     :   :
57     :   :
58     :   :
59     :   :
60     :   :
61     :   :
62     :   :
63     :   :
64     :   :
65     :   :
66     :   :
67     :   :
68     :   :
69     :   :
70     :   :
71     :   :
72     :   :
73     :   :
74     :   :
75     :   :
76     :   :
77     :   :
78     :   :
79     :   :
80     :   :
81     :   :
82     :   :
83     :   :
84     :   :
85     :   :
86     :   :
87     :   :
88     :   :
89     :   :
90     :   :
91     :   :
92     :   :
93     :   :
94     :   :
95     :   :
96     :   :
97     :   :
98     :   :
99     :   :
100    :   :
101    :   :
102    :   :
103    :   :
104    :   :
105    :   :
106    :   :
107    :   :
108    :   :
109    :   :
110    :   :
111    :   :
112    :   :
113    :   :
114    :   :
115    :   :
116    :   :
117    :   :
118    :   :
119    :   :
120    :   :
121    :   :
122    :   :
123    :   :
124    :   :
125    :   :
126    :   :
127    :   :
128    :   :
129    :   :
130    :   :
131    :   :
132    :   :
133    :   :
134    :   :
135    :   :
136    :   :
137    :   :
138    :   :
139    :   :
140    :   :
141    :   :
142    :   :
143    :   :
144    :   :
145    :   :
146    :   :
147    :   :
148    :   :
149    :   :
150    :   :
151    :   :
152    :   :
153    :   :
154    :   :
155    :   :
156    :   :
157    :   :
158    :   :
159    :   :
160    :   :
161    :   :
162    :   :
163    :   :
164    :   :
165    :   :
166    :   :
167    :   :
168    :   :
169    :   :
170    :   :
171    :   :
172    :   :
173    :   :
174    :   :
175    :   :
176    :   :
177    :   :
178    :   :
179    :   :
180    :   :
181    :   :
182    :   :
183    :   :
184    :   :
185    :   :
186    :   :
187    :   :
188    :   :
189    :   :
190    :   :
191    :   :
192    :   :
193    :   :
194    :   :
195    :   :
196    :   :
197    :   :
198    :   :
199    :   :
200    :   :
201    :   :
202    :   :
203    :   :
204    :   :
205    :   :
206    :   :
207    :   :
208    :   :
209    :   :
210    :   :
211    :   :
212    :   :
213    :   :
214    :   :
215    :   :
216    :   :
217    :   :
218    :   :
219    :   :
220    :   :
221    :   :
222    :   :
223    :   :
224    :   :
225    :   :
226    :   :
227    :   :
228    :   :
229    :   :
230    :   :
231    :   :
232    :   :
233    :   :
234    :   :
235    :   :
236    :   :
237    :   :
238    :   :
239    :   :
240    :   :
241    :   :
242    :   :
243    :   :
244    :   :
245    :   :
246    :   :
247    :   :
248    :   :
249    :   :
250    :   :
251    :   :
252    :   :
253    :   :
254    :   :
255    :   :
256    :   :
257    :   :
258    :   :
259    :   :
260    :   :
261    :   :
262    :   :
263    :   :
264    :   :
265    :   :
266    :   :
267    :   :
268    :   :
269    :   :
270    :   :
271    :   :
272    :   :
273    :   :
274    :   :
275    :   :
276    :   :
277    :   :
278    :   :
279    :   :
280    :   :
281    :   :
282    :   :
283    :   :
284    :   :
285    :   :
286    :   :
287    :   :
288    :   :
289    :   :
290    :   :
291    :   :
292    :   :
293    :   :
294    :   :
295    :   :
296    :   :
297    :   :
298    :   :
299    :   :
300    :   :
301    :   :
302    :   :
303    :   :
304    :   :
305    :   :
306    :   :
307    :   :
308    :   :
309    :   :
310    :   :
311    :   :
312    :   :
313    :   :
314    :   :
315    :   :
316    :   :
317    :   :
318    :   :
319    :   :
320    :   :
321    :   :
322    :   :
323    :   :
324    :   :
325    :   :
326    :   :
327    :   :
328    :   :
329    :   :
330    :   :
331    :   :
332    :   :
333    :   :
334    :   :
335    :   :
336    :   :
337    :   :
338    :   :
339    :   :
340    :   :
341    :   :
342    :   :
343    :   :
344    :   :
345    :   :
346    :   :
347    :   :
348    :   :
349    :   :
350    :   :
351    :   :
352    :   :
353    :   :
354    :   :
355    :   :
356    :   :
357    :   :
358    :   :
359    :   :
360    :   :
361    :   :
362    :   :
363    :   :
364    :   :
365    :   :
366    :   :
367    :   :
368    :   :
369    :   :
370    :   :
371    :   :
372    :   :
373    :   :
374    :   :
375    :   :
376    :   :
377    :   :
378    :   :
379    :   :
380    :   :
381    :   :
382    :   :
383    :   :
384    :   :
385    :   :
386    :   :
387    :   :
388    :   :
389    :   :
390    :   :
391    :   :
392    :   :
393    :   :
394    :   :
395    :   :
396    :   :
397    :   :
398    :   :
399    :   :
400    :   :
401    :   :
402    :   :
403    :   :
404    :   :
405    :   :
406    :   :
407    :   :
408    :   :
409    :   :
410    :   :
411    :   :
412    :   :
413    :   :
414    :   :
415    :   :
416    :   :
417    :   :
418    :   :
419    :   :
420    :   :
421    :   :
422    :   :
423    :   :
424    :   :
425    :   :
426    :   :
427    :   :
428    :   :
429    :   :
430    :   :
431    :   :
432    :   :
433    :   :
434    :   :
435    :   :
436    :   :
437    :   :
438    :   :
439    :   :
440    :   :
441    :   :
442    :   :
443    :   :
444    :   :
445    :   :
446    :   :
447    :   :
448    :   :
449    :   :
450    :   :
451    :   :
452    :   :
453    :   :
454    :   :
455    :   :
456    :   :
457    :   :
458    :   :
459    :   :
460    :   :
461    :   :
462    :   :
463    :   :
464    :   :
465    :   :
466    :   :
467    :   :
468    :   :
469    :   :
470    :   :
471    :   :
472    :   :
473    :   :
474    :   :
475    :   :
476    :   :
477    :   :
478    :   :
479    :   :
480    :   :
481    :   :
482    :   :
483    :   :
484    :   :
485    :   :
486    :   :
487    :   :
488    :   :

```

Query Match	60.68;	Score 129.6;	DB 5;	Length 970;
Best Local Similarity	79.08;	Pred. No. 7.2e-21;		
Matches 169; Conservative	0;	Mismatches 39;	Indels 6;	Gaps 1

QY	1	TCAAGGAGAAAGCCCTTACAGATCTGGGCCCCGGGGGTCGTGCAAGGGTGAACAAGAAAGCTGG	60
Db	26	TCAAGGAGAAAGCCCTTACAGACTGGGCCCGCCGTCCCTTTAAGGTCAAGAAGAGCTCG	85
QY	61	GCCGTGGGGGCAACCAATCCGGCTCAGGTGACCAACCAACCCCTCAACCGGAGAACGTCG	120
Db	86	GCGCTCATGTCTGGCGAAGCC-----CATCAAGTGTGTGAAGCTGAACCAAGAACGTCG	139
QY	121	TGCCCAACCATCGATGTAAGTCTGGTGGCGCTTGCACAGAGGGCCAGAACCAACGATGACCGCCCCG	180
Db	140	TGGCCCAACCATTCAAATATCTGGTCTCCGCTTTGGACAGAGGGTCAAGAACCAACGATGACCGTCCGG	199
QY	181	GGCGGCTCGAGGTGCGCGGTGGATGTGGACACAACT	214
Db	200	GGCGGCTCGAGGTGCGCGGTGGAAACCGACGACAACT	233

```

RESULT 3
US-08-313-185-57
: Sequence 57, Application US/08313185
: Patent No. 5851763
: GENERAL INFORMATION:

```

```

1  APPLICANT: Cole, Stewart
2  APPLICANT: Young, Douglas
3  APPLICANT: Zhang, Ying
4  APPLICANT: Honore, Nadine
5  APPLICANT: Telenti, Amalio
6  APPLICANT: Bodmer, Thomas
7  TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
8  TITLE OF INVENTION: In Mycobacterium Tuberculosis
9  NUMBER OF SEQUENCES: 66
10 CORRESPONDENCE ADDRESSES:
11 ADDRESSSEE: Finnegan, Henderson, Farabow, Garrett &
12 ADDRESSSEE: Dunner
13 STREET: 1300 I Street, N.W.
14 CITY: Washington
15 STATE: D.C.
16 COUNTRY: USA
17 ZIP: 20005-3315
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: Patentin Release #1.0, Version #1.25
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/08/313,185
25 FILING DATE: 12-OCT-1994
26 CLASSIFICATION: 435
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Meyers, Kenneth J.
29 REGISTRATION NUMBER: 25,146
30 REFERENCE/DOCKET NUMBER: 02356, 0068-00000
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: (202) 408-4000
33 TELEFAX: (202) 408-4400
34 INFORMATION FOR SEQ ID NO: 57:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 3447 base pairs
37 TYPE: nucleic acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 MOLECULE TYPE: DNA (genomic)
41 US-08-313-185-57

```

Query Match	50.8%;	Score 108.8;	DB 2;	Length 3447;
Best Local Similarity	72.9%;	Pred. No. 3.1e-16;		
Matches 156;	Conservative	0;	Mismatches 52;	Indels 6;
				Gaps 1;

QY	1	TCAAGAGAAAGCGCTACAGATCTGGGCGCCGCGGGTGGTAAAGAGTGAACAAAGAAAGCTGG	60
Db	809	TCAAGAGAAAGCGCTACAGACTGGCCAGGGTGGTGTTCACAAAGTCAAGAAAGCTCG	868
QY	61	GCCGTGGGGGCAACCAATCCGGCTCAGGTGACCCACCAACCTTCACCGAGGAAAGCTGG	120
Db	869	GGTTGGACCCCGGTGAGTTGA-----TCACGTGCTCACGCTGACCGAAGAGATGTGG	922
QY	121	TGCCCAACATTCAGTACTGTGGTGGCTGCACAGAGGGCCAGACCAAGATGACCGCCCG	180
Db	923	TGCCCAACATTAAGTACTGGTTGTTCGTGATAGGGTCAGTGCACAAATGACTGTCCAG	982
QY	181	GCGGCTCAGAGTCCCGGTGGATGTGGACACAT	214
Db	983	GTGGGTTACAACTGCCAATGGAACCTACGATAT	1016

RESULT 4
US-09-082-614A-57

Sequence 57, Application US/09082614A
Patent No. 6124098
GENERAL INFORMATION:
APPLICANT: Heym, Beate
APPLICANT: Cole, Stewart
APPLICANT: Young, Douglas
APPLICANT: Zhang, Ying
APPLICANT: Honore, Nadine
APPLICANT: Telenti, Amelio
APPLICANT: Bodmer, Thomas
TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
TITLE OF INVENTION: In Mycobacterium Tuberculosis
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,614A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/313,185
FILING DATE: 12-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 02356.0068-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4000
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 3447 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-082-614A-57

Query Match 50.8%; Score 108.8; DB 3; Length 3447;
Best Local Similarity 72.9%; Pred. No. 3.1e-16;
Matches 156; Conservative 0; Mismatches 52; Indels 6; Gaps 1;

QY 1 TCAGAGAGAGCGCTACGATCGTGGCCCGCGTGGTGGTACAGGTGAACAAGAGCTGG 60
DB 809 TCAGAGAGAGCGCTACGATCGTGGCCAGGTTGTCGTACAGGTCAACAAGAGCTGG 868
QY 61 GCCTGGCGCGGACCAATCCGCTCAAGTGACACACCACTCAACGAGAGAGCTGG 120
DB 869 GGTTCACACCGCGGTAGTTGA-----TCACGTCGTCCTCCACGCTGACGAGAGAGATGTCG 922
QY 121 TGCACACCATGACGTACCTGCTGCTGCACAGAGGCGCAGACAGCATGACCGCCCGCCG 180
DB 923 TGCACACCATGACGTACCTGCTGCTGCATGAGGTCACATGCAATGACTGTCCAG 982
QY 181 GGGGGGTGAGGTGCGGTGATGTGACGACAT 214
DB 983 GTGGGTGAGAGTCCAGTGAAGACTGACGATAT 1016

RESULT 5
5352575-8
Patent No. 5352575

APPLICANT: PETROVSKIS, ERIK A.; POST, LEONARD E.; TIMMINS, JAMES G.
TITLE OF INVENTION: PSEUDORABIES VIRUS PROTEIN
NUMBER OF SEQUENCES: 12
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/513,282
FILING DATE: 20-APR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 100,817
FILING DATE: 29-JUN-1987
APPLICATION NUMBER: 886,260
FILING DATE: 16-JUL-1986
APPLICATION NUMBER: 784,787
FILING DATE: 04-OCT-1985
APPLICATION NUMBER: 801,799
FILING DATE: 26-NOV-1985
APPLICATION NUMBER: 844,113
FILING DATE: 26-MAR-1986
SEQ ID NO: 8
LENGTH: 1734
5352575-8

Query Match 19.7%; Score 42.2; DB 6; Length 1734;
Best Local Similarity 54.18; Pred. No. 0.18;
Matches 86; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 56 GCTGGCCCTGGCGCGCACCAATCCGCTCAGTGACACACCACTCAGGAGAGAGA 115
DB 1065 GATGACCCACAAAGCGCCAGCTGCGCACTGAGACTGACAGCTGTGTCACCGCGCGGA 1124
QY 116 CGTGTGCGCACCATGATGACTGTGCGCTGCACGAGGCGCAGACGATGACCGC 175
DB 1125 GTACGTACGCTGATCAAGAGCTGACGCGCCGCGCGCGGCGGAGCACCCTGGGG 1184
QY 176 CCCCGCGCGCTGAGGTGCGGTGATGTGACGACAT 214
DB 1185 CCCCGCGCGCGGAGAGACGCGATCTACGTGACGCGCT 1223

RESULT 6
US-09-335-409-1
Sequence 1, Application US/09335409
Patent No. 6121029
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/335,409
CURRENT FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 68750
TYPE: DNA
ORGANISM: Sorangium cellulosum
US-09-335-409-1

Query Match 17.9%; Score 38.4; DB 3; Length 68750;
Best Local Similarity 52.58; Pred. No. 1.5;
Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 19 ATCTGCCCCGCTGGTCTGTGTAAGGTGAACAAGAGCTGGCGCTGGCGCACCAATC 78
DB 21325 ATCTGCGGAGCGCGCGCGCGCGCACCGGAGCGGAGACTGTCTCCGCGCTGGCGCG 21384
QY 79 CGGCTCAGGTGACCAACCAACCACTCAACGAGAGAGCTGTGCGCACCATGAGTACC 138
DB 21385 CGAGCGGAGCGCGGAGCGCGCTCTCGGAGCGCGCTCTCGGAGAGTCTCGACAG 21444

CURRENT APPLICATION NUMBER: US/09/568,486
CURRENT FILING DATE: 2009-05-10

;; PRIOR APPLICATION NUMBER: 09/335,409
;; PRIOR FILING DATE: 1999-06-17
;; NUMBER OF SEQ ID NOS: 30
;; SOFTWARE: Patent In Ver. 2.0
;; SEQ ID NO 1
;; LENGTH: 68750
;; TYPE: DNA
;; ORGANISM: Sorangium cellulosum
US-09-568-486-1

Query Match 17.9%: Score 38.4; DB 4; Length 68750;
Best Local Similarity 52.5%; Pred. No. 1.5;
Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 19 ATCTGCCCCGCTGGTGTGCTACAGAGTGAACAAGAGCTGGGCGGCGCAATC 78
DB 21325 ATCGCGGAGCG 21384
OY 79 CGGCTCAGCTGACCAACCACTTACACGAGAGAGCTGTGCGCCACCATGAGTACC 138
DB 21385 CCGAGCCGAGCGCGCGCGCGCGCGCTCTCGGAGCGCGCTCTCGCGCGAGATCTCGCAGG 21444
OY 139 TGTGCGCGCTGCAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 178
DB 21445 TGTGCGCGCTGCAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 21484

RESULT 11
US-09-568-472-1

;; Sequence 1, Application US/09568472

;; Patent No. 6358719

;; GENERAL INFORMATION:

;; APPLICANT: Schupp, Thomas

;; APPLICANT: Ligon, James

;; APPLICANT: Molnar, Istvan

;; APPLICANT: Zirkle, Ross

;; APPLICANT: Cyr, Devon

;; APPLICANT: Goerlach, Joern

;; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES

;; FILE REFERENCE: 4-30582A

;; CURRENT APPLICATION NUMBER: US/09/568,472

;; CURRENT FILING DATE: 2000-05-10

;; PRIOR APPLICATION NUMBER: 09/335,409

;; PRIOR FILING DATE: 1999-06-17

;; NUMBER OF SEQ ID NOS: 30

;; SOFTWARE: Patent In Ver. 2.0

;; SEQ ID NO 1

;; LENGTH: 68750

;; TYPE: DNA

;; ORGANISM: Sorangium cellulosum

US-09-568-472-1

Query Match 17.9%: Score 38.4; DB 4; Length 68750;
Best Local Similarity 52.5%; Pred. No. 1.5;
Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 19 ATCTGCCCCGCTGGTGTGCTACAGAGTGAACAAGAGCTGGGCGGCGCAATC 78
DB 21325 ATCGCGGAGCG 21384
OY 79 CGGCTCAGCTGACCAACCACTTACACGAGAGAGCTGTGCGCCACCATGAGTACC 138
DB 21385 CCGAGCCGAGCGCGCGCGCGCGCGCTCTCGGAGCGCGCTCTCGCGCGAGATCTCGCAGG 21444
OY 139 TGTGCGCGCTGCAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 178
DB 21445 TGTGCGCGCTGCAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 21484

RESULT 12
US-09-567-899-1
;; Sequence 1, Application US/09567899
;; Patent No. 6383787

;; GENERAL INFORMATION:
;; APPLICANT: Schupp, Thomas
;; APPLICANT: Ligon, James
;; APPLICANT: Molnar, Istvan
;; APPLICANT: Zirkle, Ross
;; APPLICANT: Cyr, Devon
;; APPLICANT: Goerlach, Joern
;; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
;; FILE REFERENCE: 4-30582A
;; CURRENT APPLICATION NUMBER: US/09/567,899
;; CURRENT FILING DATE: 2000-05-10
;; PRIOR APPLICATION NUMBER: 09/335,409
;; PRIOR FILING DATE: 1999-06-17
;; NUMBER OF SEQ ID NOS: 30
;; SOFTWARE: Patent In Ver. 2.0
;; SEQ ID NO 1
;; LENGTH: 68750
;; TYPE: DNA
;; ORGANISM: Sorangium cellulosum
US-09-567-899-1

Query Match 17.9%: Score 38.4; DB 4; Length 68750;
Best Local Similarity 52.5%; Pred. No. 1.5;
Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

OY 19 ATCTGCCCCGCTGGTGTGCTACAGAGTGAACAAGAGCTGGGCGGCGCAATC 78
DB 21325 ATCGCGGAGCG 21384
OY 79 CGGCTCAGCTGACCAACCACTTACACGAGAGAGCTGTGCGCCACCATGAGTACC 138
DB 21385 CCGAGCCGAGCGCGCGCGCGCGCGCTCTCGGAGCGCGCTCTCGCGCGAGATCTCGCAGG 21444
OY 139 TGTGCGCGCTGCAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 178
DB 21445 TGTGCGCGCTGCAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 21484

RESULT 13

US-08-881-784-8

;; Sequence 8, Application US/08881784

;; Patent No. 6083731

;; GENERAL INFORMATION:

;; APPLICANT: Croteau, Rodney B.

;; APPLICANT: Lupien, Shari L.

;; APPLICANT: Karp, Frank

;; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR

;; NUMBER OF SEQUENCES: 58

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness

;; STREET: 1420 Fifth Avenue, Suite 2800

;; CITY: Seattle

;; STATE: WA

;; COUNTRY: USA

;; ZIP: 98101

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; OPERATING SYSTEM: IBM PC compatible

;; SOFTWARE: Patent Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/881,784

;; FILING DATE:

;; CLASSIFICATION: 435

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Shelton, Dennis K.

;; REGISTRATION NUMBER: 26,997

;; REFERENCE/DOCKET NUMBER: WSUR19777

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: (206) 224-0718

;; TELEFAX: (206) 224-0779

INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1665 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Mentha x piperita
IMMEDIATE SOURCE:
CLONE: PM17
US-08-881-784-8

Query Match
Best Local Similarity 51.8%; Score 38; DB 3; Length 1665;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 23 GCGCCGCGTGGTGGTCAAGGTGAACAGAGCTGGGCGGCGGCAATCCGGC 82
DB 270 GTCCCGCGAGGCGGAGGAGGCGGATGACCTGTGGACCGGCGCGGCGGTT 329
QY 83 TCAGGTGACACACACCTCTACCGAGAGAGCTGGCCACCTCGAGTACTGGT 142
DB 330 CGAGAGCATCGGAGGAGAGATCATGTGTACGACACGACGACTCTTCAGCCCTA 389
QY 143 GCGCTGACAGAGGCGGACGACGATGACCGCCCGCGCGCGCTC 188
DB 390 CAGCGTGCACCTGGCGCGGAGATGCGGAGATCTCGCTCCGAGCTC 435

RESULT 14

US-09-292-768-3
Sequence 3, Application US/09292768
Patent No. 6194185
GENERAL INFORMATION:
APPLICANT: Croteau, Rodney B
APPLICANT: Lupien, Shari L
APPLICANT: Karp, Frank
TITLE OF INVENTION: LIMONENE HYDROXYLASES
TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
FILE REFERENCE: wsr13463
CURRENT APPLICATION NUMBER: US/09/292,768
CURRENT FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 08/881,784
EARLIER FILING DATE: 1997-06-24
NUMBER OF SEQ ID NOS: 70
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3
LENGTH: 1665
TYPE: DNA
ORGANISM: Mentha piperita
FEATURE:
NAME/KEY: CDS
LOCATION: (19)..(1518)
US-09-292-768-3

Query Match
Best Local Similarity 17.8%; Score 38; DB 4; Length 1665;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 23 GCGCCGCGTGGTGGTCAAGGTGAACAGAGCTGGGCGGCGGCAATCCGGC 82
DB 270 GTCCCGCGAGGCGGAGGAGGCGGATGACCTGTGGACCGGCGCGGCGGTT 329
QY 83 TCAGGTGACACACACCTCTACCGAGAGAGCTGGCCACCTCGAGTACTGGT 142
DB 330 CGAGAGCATCGGAGGAGAGATCATGTGTACGACACGACGACATCATTCAGCCCTA 389
QY 143 GCGCTGACAGAGGCGGACGACGATGACCGCCCGCGCGCGCTC 188
DB 390 CAGCGTGCACCTGGCGCGGAGATGCGGAGATCTCGCTCCGAGCTC 435

RESULT 15
US-09-292-768-67
Sequence 67, Application US/09292768
Patent No. 6194185
GENERAL INFORMATION:
APPLICANT: Croteau, Rodney B
APPLICANT: Lupien, Shari L
APPLICANT: Karp, Frank
TITLE OF INVENTION: LIMONENE HYDROXYLASES
TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
FILE REFERENCE: wsr13463
CURRENT APPLICATION NUMBER: US/09/292,768
CURRENT FILING DATE: 1999-04-14
EARLIER APPLICATION NUMBER: 08/881,784
EARLIER FILING DATE: 1997-06-24
NUMBER OF SEQ ID NOS: 70
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 67
LENGTH: 1665
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: Computer-generated nucleic acid sequence encoding
OTHER INFORMATION: Limonene-3-hydroxylase
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)..(1665)
OTHER INFORMATION: Computer-generated nucleic acid sequence encoding
FEATURE:
NAME/KEY: CDS
LOCATION: (19)..(1518)
US-09-292-768-67

Query Match
Best Local Similarity 17.8%; Score 38; DB 4; Length 1665;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 23 GCGCCGCGTGGTGGTCAAGGTGAACAGAGCTGGGCGGCGGCAATCCGGC 82
DB 270 GTCCCGCGAGGCGGAGGAGGCGGATGACCTGTGGACCGGCGCGGCGGTT 329
QY 83 TCAGGTGACACACACCTCTACCGAGAGAGCTGGCCACCTCGAGTACTGGT 142
DB 330 CGAGAGCATCGGAGGAGAGATCATGTGTACGACACGACGACATCATTCAGCCCTA 389
QY 143 GCGCTGACAGAGGCGGACGACGATGACCGCCCGCGCGCGCTC 188
DB 390 CAGCGTGCACCTGGCGCGGAGATGCGGAGATCTCGCTCCGAGCTC 435

Search completed: November 12, 2002, 21:03:28
Job time : 236.698 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	214	100.0	214	22	AA505221	
2	186.8	87.3	214	22	AA505207	Mycobacterium abscessus
3	153.6	71.8	208	22	AA505219	Mycobacterium chelonae
4	145.6	68.0	208	22	AA505215	Mycobacterium fortuitum
5	145.2	67.9	214	22	AA505214	Mycobacterium avium
6	142.4	66.5	208	22	AA505201	Mycobacterium malvarum
7	142.4	66.5	208	22	AA505206	Mycobacterium goodii
8	142.4	66.5	208	22	AA505210	Mycobacterium kansasii
9	142	66.4	211	22	AA505218	Mycobacterium ulcerans

10	140.8	65.8	208	22	AAOS5211
11	140.8	65.8	208	22	AAOS5211
12	137.6	64.3	208	22	AAOS5207
13	136	63.6	208	22	AAOS5220
14	136	63.6	208	22	AAOS5220
15	129.6	60.6	208	22	AAOS5205
16	129.6	60.6	208	22	AAOS5200
17	129.6	60.6	208	22	AAOS5216
18	129.6	60.6	970	17	AAOT09670
19	129.6	60.6	3519	22	AAH51976
20	129.6	60.6	3534	22	AAH02079
21	129.6	60.6	3853	21	AAAT746515
22	129.6	60.6	3853	21	AAAB899292
23	128	59.8	208	22	AAOS5222
24	128	59.8	208	22	AAOS5222
25	121.6	56.8	205	22	AAOS5220
26	120.8	56.4	205	22	AAOS5212
27	119.2	55.7	207	22	AAOS5213
28	108.8	50.8	3447	14	AAOS155303
29	107.8	50.4	223	22	AAOS5213
30	103	48.1	223	22	AAOS5209
31	80.4	37.6	3395	22	AAH65512
32	80.4	37.6	349890	22	AAH68587
33	55.2	25.8	27426	22	AAH59541
34	43.8	20.5	1734	21	AAAH0376
35	43.8	20.5	5895	23	ABN87056
36	42.8	20.2	1844	21	AAO12000
37	42.8	20.0	77536	21	AAO14651
38	42.2	19.7	1734	8	AAAT7095
39	42.2	19.7	1734	15	AAO73450
40	42.2	19.7	1734	22	AAOS098322
41	42.2	19.3	1734	22	AAPO3607
42	41.4	19.3	547	21	AAAC03937
43	40.4	18.9	10919	22	AAAS08653
44	40.2	18.8	537	21	AAAC56469
45	39.6	18.5	2853	23	ABLI2631

ALIGNMENTS

RESULT 1	AA050221	ID	AA050221 standard; DNA, 214 BP.
XX	AA050221:		
AC	07-SEP-2001	(first entry)	
DT			
XX			
DE	Mycobacterium abscessus	rpob gene fragment.	
XX			
KM	Non-tuberculous mycobacteria: rpob gene fragment; NTM; HIV; FRA; RFLP		
KW	PCR-restriction fragment length polymorphism analysis; ds.		
XX			
OS	Mycobacterium abscessus.		
PN	W0200131061-A1.		
XX			
PD	03-MAY-2001.		
XX			
PF	27-OCT-2000; 2000WO-KR01223.		
XX			
PR	27-OCT-1999; 99KR-0046795.		
XX			
PA	(ERUM-) ERUME BIOTECH CO LTD.		
XX			
PI	Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;		
XX			
DR	WPI: 2001-300520/31.		
XX			
PT	New DNA fragments from the rpob gene of mycobacteria, useful for		
PT	diagnosis and identification of many mycobacterial species by		
PT	restriction fragment length polymorphism -		

XX
PS
Claim 1; Page 46; 50pp; English.

The present sequence for *Mycobacterium abscessus* rpoB gene fragment is 1 of 24 rpoB gene fragments (AA505201-AA505224) from various *Mycobacterial* species. These rpoB gene fragments can be used in the diagnosis and identification of *Mycobacterium* species using a novel PCR-restriction fragment length polymorphism analysis (PRA) method. The method comprises obtaining a restriction fragment length polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA fragment from the microorganism to be identified and comparing the RFLP patterns from the known rpoB gene fragments with the unidentified fragment. The rpoB gene fragments are useful to identify a wide range of *Mycobacterium* species, e.g. for diagnosis or to obtain epidemiological and pathogenesis information for selection of appropriate therapies, including *M. tuberculosis*, *M. leprae* and non-tuberculous *Mycobacteria* (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and cost effective (only 1 PCR required), and can differentiate between many species in a single experiment, including those difficult to distinguish by usual biochemical tests. Also described are oligonucleotide probes (AA505227-AA505242) for detecting specific *Mycobacterial* species.

SQ Sequence 214 BP; 45 A; 69 C; 72 G; 28 T; 0 other;

Query Match	100.08;	Score 214;	DB 22;	Length 214;
Best Local Similarity	100.08;	Pred. No. 5.4e-38;		
Matches 214: Conservative	0;	Mismatches	0;	Indels

QY	1	TCMAAGAGAAAGGCGTACACATCTGGCGCCGCTCGGTACAAAGGTGAACAAAGACCTGG	60
Db	1	TCMAAGAGAAAGGCGTACATCTGGCGCCCGCTCGGTACAAAGGTGAACAAAGACCTGG	60
QY	61	GCCTGGCGGCGACCAATCTCGGCTGAGGTGATGCACACACACCTTCACCGAGGAAGACGTGG	12
Db	61	GCCTGGCGGCGACCAATCTCGGCTGAGGTGATGCACACACACCTTCACCGAGGAAGACGTGG	12
QY	121	TCGCCACCATCGATGACTTGGTGGCGCTGCACGAGGGCCAGACACAGATGACCGCCCGG	18
Db	121	TCGCCACCATCGAGTACTTGGTGGCGCTGCACGAGGGCCAGACACAGATGACCGCCCGG	18
QY	181	GGGGGCTGAGAGTGGCGGCTGGATGTGGACGACAT	214
Db	181	GGGGGCTGAGAGTGGCGGCTGGATGTGGACGACAT	214

RESULT 2
AA505707

ID	standard; DNA; 214 BP.
AAS05207	

AC AAS05207;

DT 07-SEP-2001 (first entry)

Mycobacterium chelonae rpoB gene fragment.

Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP; PCR-restriction fragment length polymorphism analysis; ds.

Mycobacterium chelonae.

PN WO200131061-A1.

PD 03-MAY-2001.

27-OCT-2000; 2000WO-KR01223.

27-OCT-1999; 99KR-0046795.

(ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.
DR

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
 PT diagnosis and identification of many mycobacterial species by
 PT restriction fragment length polymorphism -
 XX
 PS Claim 1, Page 42: 50pp; English.

The present sequence for *Mycobacterium chelonae* rpoB gene fragment is 1 of 24 rpoB gene fragments (AA505201-AA505224) from various *Mycobacterium* species. These rpoB gene fragments can be used in the diagnosis and identification of *Mycobacterium* species using a novel PCR-restriction fragment length polymorphism analysis (PRA) method. The method comprises obtaining a restriction fragment length polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA fragment from the microorganism to be identified and comparing the RFLP patterns from the known rpoB gene fragments with the unidentified fragment. The rpoB gene fragments are useful to identify a wide range of *Mycobacterium* species, e.g. for diagnosis or to obtain epidemiological and pathogenesis information for selection of appropriate therapies, including *M. tuberculosis*, *M. leprae* and non-tuberculous mycobacteria (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and cost effective (only 1 PCR required), and can differentiate between many species in a single experiment, including those difficult to distinguish by usual biochemical tests. Also described are oligonucleotide probes (AA505227-AA505242) for detecting specific *Mycobacterium* species.

SQ Sequence 214 BP; 41 A; 72 C; 72 G; 29 T; 0 other;

Query Match	87.38;	Score 186.8;	DB 22;	Length 214;
Best Local Similarity	92.18;	Pred. No. 4.6e-32;		
Matches 197; Conservative	0;	Mismatches 17;		

QY	1	TCAGAGAGAGCGCTACGATCTGGCCCGCGCTGGGTGCATCAAGGTAAGAAAGAAAGCTGG	60
Db	1	TCAGAGAGAAAGCGTACGACTGGCCCGCGGTGGCGGTACAAAGGTGAACAAAGAAAGCTGG	60
QY	61	GCCTGGCGGGACCAATCCGGCTAGGTGACCAACCAACCTTCAACCGAGGAAGACTGG	120
Db	61	GTCTTTGGCGGTGCACACCCGGCTCTGGTGACTCCCAACAGCTTCACCGAAGAAACGCTGG	120
QY	121	TGCGCACCATGAGTACTGTTGGCCTCGACAGAGGGCCAGACCAAGATGAGCCGCCCCCG	180
Db	121	TGCGCACCATGAGGTAACCTGGTGGCCCTGACAGAGGGCCAGACCAAGATGAGCCGCCCCCG	180
QY	181	GGCGGCTGAGAGTCCGGTGTGATGTGAGACACAT	214
Db	181	GGCGGCTGAGAGTCCGGTGTGAGTGTGAGACACAT	214

RESULT 3
AAS05219

ID AAS05219 standard; DNA; 208 BP.

AC AAS05219;

DT	07-SEP-2001	(first entry)
XY		

Mycobacterium fortuitum rpoB gene fragment.

non-tuberculous mycobacteria; ipob gene fragment; NTM; HIV; PRA; RFLP; PCR-restriction fragment length polymorphism analysis; ds.

Mycobacterium fortuitum.

WO200131061-A1

03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

```
XX 27-OCT-1999; 99KR-0046795.
PR (ERUM-) ERMIE BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI; 2001-300520/31.
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism
XX
XX Claim 1: Page 46; 50pp; English.
XX
XX The present sequence for Mycobacterium fortuitum rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterial species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments: isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. lepre
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 43 A; 69 C; 69 G; 27 T; 0 other:
Query Match 71.8%; Score 153.6; DB 22; Length 208;
Best Local Similarity 86.0%; Pred. No. 7.8e-25;
Matches 184; Conservative 0; Mismatches 24; Indels 6; Gaps 1;
OY 1 TCAAGAGAACGCGTACGATCTGCGCGGTCGCTACAGCGTGAACAGAACCTGG 60
DB 1 TCAAGAGAACGCGTACGATCTGCGCGGTCGCTACAGCGTGAACAGAACCTGG 60
OY 61 GCTGGGCGGACCAATCCGGCTCAGGTGACACCAACCACTCACCAGGAAGACGTG 120
DB 61 GCTGGGCGGACCAATCCGGCTCAGGTGACACCAACCACTCACCAGGAAGACGTG 120
OY 121 TCGGCACCATCGATAGTACCTGCTGCGCTGACAGAGGCCACACAGATGACCGCCCG 180
DB 115 TCGGCACCATCGATAGTACCTGCTGCGCTGACAGAGGCCACACAGATGACCGCCCG 174
OY 181 GCGGCGTGAAGTCCGCGTGGATGTGAGACACAT 214
DB 175 GCGGCGTGAAGTCCGCGTGGATGTGAGACACAT 208
RESULT 4
AAS05215
ID AAS05215 standard; DNA: 208 BP.
XX
XX AAS05215;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium avium rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium avium.
```

```
XX
XX MO200131061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERMIE BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism
XX
XX Claim 1: Page 44; 50pp; English.
XX
XX The present sequence for Mycobacterium avium rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments: isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. lepre
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other:
Query Match 68.0%; Score 145.6; DB 22; Length 208;
Best Local Similarity 83.6%; Pred. No. 4.3e-23;
Matches 179; Conservative 0; Mismatches 29; Indels 6; Gaps 1;
OY 1 TCAAGAGAACGCGTACGATCTGCGCGGTCGCTACAGCGTGAACAGAACCTGG 60
DB 1 TCAAGAGAACGCGTACGATCTGCGCGGTCGCTACAGCGTGAACAGAACCTGG 60
OY 61 GCTGGGCGGACCAATCCGGCTCAGGTGACACCAACCACTCACCAGGAAGACGTG 120
DB 61 GCTGGGCGGACCAATCCGGCTCAGGTGACACCAACCACTCACCAGGAAGACGTG 120
OY 121 TCGGCACCATCGATAGTACCTGCTGCGCTGACAGAGGCCACACAGATGACCGCCCG 180
DB 115 TCGGCACCATCGATAGTACCTGCTGCGCTGACAGAGGCCACACAGATGACCGCCCG 174
OY 181 GCGGCGTGAAGTCCGCGTGGATGTGAGACACAT 214
DB 175 GCGGCGTGAAGTCCGCGTGGATGTGAGACACAT 208
RESULT 5
AAS05214
ID AAS05214 standard; DNA: 214 BP.
XX
XX AAS05214;
XX
XX 07-SEP-2001 (first entry)
XX
```

DE Mycobacterium malmoeense rpoB gene fragment.

XX Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.
XX

OS Mycobacterium malmoeense.

XX W0200131061-A1.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000MO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI: 2001-300520/31.

XX

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism

XX Claim 1; Page 44; 50pp; English.

XX The present sequence for Mycobacterium malmoeense rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC selection of appropriate epidemiological and pathogenesis information for
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC experiment), and can differentiate between many species in a single
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

XX Sequence 214 BP; 42 A; 68 C; 75 G; 29 T; 0 other;

Query Match 67.9%; Score 145.2; DB 22; Length 214;
Best Local Similarity 79.9%; Pred. No. 5.3e-23;
Matches 171; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1 TCAAGAGAGAGCGCTACGATCTGCGCCCGCTGGTGGTACAGAGTGAACAAGAGCTG 60
DB 1 TCAAGAGAGAGCGCTACGATCTGCGCCCGCTGGTGGTGGTACAGAGTGAACAAGAGCTG 60

QY 61 GCGTGGGCGGACCAATCGGCTCAGGTGACCAACACACCTTCACCGAGGAAGCTGC 120
DB 61 GCGTGGGCGGCGGACCAATCGGCTCAGGTGACCAACACACCTTCACCGAGGAAGCTGC 120

QY 121 TCGCCACCATGAGTACCTGCTGCGCTGCAGCAGGCGCAGACAGTACGCCGCCCG 180
DB 121 TCGCCACCATGAGTACCTGCTGCGCTGCAGCAGGCGCAGACAGTACGCCGCCCG 180

QY 181 GCGGCGTGAAGTGCAGGTGATGTGACGACAT 214
DB 181 GCGGCGTGAAGTGCAGGTGATGTGACGACAT 214

RESULT 6
AAS05201

ID AAS05201 standard; DNA: 208 BP.

XX AAS05201;

XX 07-SEP-2001 (first entry)

XX Mycobacterium gordonae type I rpoB gene fragment.

XX Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.
XX

OS Mycobacterium gordonae type I.

XX W0200131061-A1.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000MO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI: 2001-300520/31.

XX

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism

XX Claim 1; Page 40; 50pp; English.

XX The present sequence for Mycobacterium gordonae type I rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC selection of appropriate epidemiological and pathogenesis information for
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC experiment), and can differentiate between many species in a single
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 46 A; 71 C; 65 G; 26 T; 0 other;

Query Match 66.5%; Score 142.4; DB 22; Length 208;
Best Local Similarity 82.7%; Pred. No. 2.1e-22;
Matches 177; Conservative 0; Mismatches 31; Indels 6; Gaps 1;

QY 1 TCAAGAGAGAGCGCTACGATCTGCGCCCGCTGGTGGTGAAGTGAACAAGAGCTG 60
DB 1 TCAAGAGAGAGCGCTACGATCTGCGCCCGCTGGTGGTGAAGTGAACAAGAGCTG 60

QY 61 GCGTGGGCGGACCAATCGGCTCAGGTGACCAACACCTTCACCGAGGAAGCTGC 120
DB 61 GCGTGGGCGGCGGACCAATCGGCTCAGGTGACCAACACCTTCACCGAGGAAGCTGC 120

QY 121 TCGCCACCATGAGTACCTGCTGCGCTGCAGCAGGCGCAGACAGTACGCCGCCCG 180
DB 121 TCGCCACCATGAGTACCTGCTGCGCTGCAGCAGGCGCAGACAGTACGCCGCCCG 180

QY 181 GCGGCGTGAAGTGCAGGTGATGTGACGACAT 214

Db 175 GCGGCACCGAGTGCCTGTGAGCAGCAGACAT 208

RESULT 7

ID AAS05208 standard; DNA: 208 BP.

XX AAS05208;

DT 07-SEP-2001 (first entry)

XX Mycobacterium kansasii rpoB gene fragment.

XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.

XX Mycobacterium kansasii.

XX W0200131061-A1.

XX 03-MAY-2001.

PF 27-OCT-2000; 2000MO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -

PS Claim 1; Page 42; 50pp; English.

XX The present sequence for Mycobacterium kansasii rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;

Query Match 66.5%; Score 142.4; DB 22: Length 208;

Best Local Similarity 82.7%; Pred. No. 2.le-22;

Matches 177; Conservative 0; Mismatches 31; Indels 6; Gaps 1;

OY 1 TCAAGGAGAGCGCTACGATCTGCGCCGCTGGTCAAGGTGAACAAGAGCTGG 60

DB 1 TCAAGGAGAGCGCTACGATCTGCGCCGCTGGTCAAGGTGAACAAGAGCTGG 60

OY 61 GCGTGGGCGCAGCAATCGGCTAGGTGAGCAGCAGCAGCAGCAGGAGAGAGCTCG 120

DB 61 GCGTGGGCGCAGCAATCGGCTAGGTGAGCAGCAGCAGCAGCAGGAGAGAGCTCG 114

OY 121 TCGCCACCATCGATCTGCTGCGCTGCAGAGGCGCCAGACCAAGATGACCGCCCG 180
DB 115 TCGCCACCATCGATCTGCTGCGCTGCAGAGGCGCCAGACCAAGATGACCGCG 174
OY 181 GCGGCGTCGAGGTGCGCGGTGAGTGTGAGCAGACAT 214
DB 175 GCGGCGTCGAGGTGCGCGGTGAGTGTGAGCAGACAT 208

RESULT 8

ID AAS05210 standard; DNA: 208 BP.

XX AAS05210;

DT 07-SEP-2001 (first entry)

XX Mycobacterium ulcerans rpoB gene fragment.

XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.

XX Mycobacterium ulcerans.

XX W0200131061-A1.

XX 03-MAY-2001.

PF 27-OCT-2000; 2000MO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -

PS Claim 1; Page 43; 50pp; English.

XX The present sequence for Mycobacterium ulcerans rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 45 A; 68 C; 67 G; 28 T; 0 other;

Query Match 66.5%; Score 142.4; DB 22: Length 208;

Best Local Similarity 82.7%; Pred. No. 2.le-22;

Matches 177; Conservative 0; Mismatches 31; Indels 6; Gaps 1;

OY 1 TCAAGGAGAGCGCTACGATCTGCGCCGCTGGTCAAGGTGAACAAGAGCTGG 60

Db 1 TCAAGGAGAACGCTACGACTGCTGCGCTGGGTCGGTACAGGCAACAGAGAGCTCG 60
QY 61 GCCTGGGGCGACCAATCCGGCTCAGGTGACACCAACCACTTACCCAGAGAGAGCTCG 120
Db 61 GCCTGAACGCGCGCCAGCC-----CATCACAGCTCGACGCTGACCCAGAGAGAGCTCG 114
QY 121 TCGCCACGATCGATGCTGCGGCTGACAGAGGCGCAGACCAATGACCGCCCGCG 180
Db 115 TCGCCACGATCGATGCTGCGGCTGACAGAGGCGCAGACCGATGACCGCTCGCG 174
QY 181 GCGGCTCGAGGTGCGGCTGATGTGACGACAT 214
Db 175 GCGGTGTGAGGTGCGGCTGACAGACGACAT 208

RESULT 9

AAS05218 ID AAS05218 standard; DNA; 211 BP.

AAS05218; AC

DT 07-SEP-2001 (first entry)

DE Mycobacterium flavescens rpoB gene fragment.

KW Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium flavescens.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism

PS Claim 1; Page 45; 50pp; English.

The present sequence for Mycobacterium flavescens rpoB gene fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from various Mycobacterium species. These rpoB gene fragments can be used in the diagnosis and identification of Mycobacterium species using a novel PCR-restriction fragment length polymorphism analysis (PRA) method. The method comprises obtaining a restriction fragment length polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA fragment from the microorganism to be identified and comparing the RFLP patterns from the known rpoB gene fragments with the unidentified fragment. The rpoB gene fragments are useful to identify a wide range of Mycobacterium species, e.g. for diagnosis or to obtain epidemiological and pathogenesis information for selection of appropriate therapies, including M. tuberculosis, M. leprae and non-tuberculous mycobacteria (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and cost effective (only 1 PCR required), and can differentiate between many species in a single experiment, including those difficult to distinguish by usual biochemical tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for detecting specific Mycobacterium species.

Sequence 211 BP; 49 A; 65 C; 70 G; 27 T; 0 other;

Query Match 66.4%; Score 142; DB 22; Length 211;
Best Local Similarity 82.2%; Pred. No. 2,66-22;
Matches 176; Conservative 0; Mismatches 35; Indels 3; Gaps 1;
QY 1 TCAAGGAGAACGCTACGATCTGCGCGGCTGAGGTGCTACAGGTGACCAAGAGAGCTCG 60
Db 1 TCAAGGAGAACGCTACGATCTGCGCGGCTGAGGTGCTACAGGTGACCAAGAGAGCTCG 60
QY 61 GCCTGGGGCGACCAATCCGGCTCAGGTGACACCAACCACTTACCCAGAGAGAGCTCG 120
Db 61 GCCTGAACGCGCGCCAGCC-----CATCACAGCTCGACGCTGACCCAGAGAGAGCTCG 114
QY 121 TCGCCACGATCGATGCTGCGGCTGACAGAGGCGCAGACCAATGACCGCCCGCG 180
Db 115 TCGCCACGATCGATGCTGCGGCTGACAGAGGCGCAGACCGATGACCGCTCGCG 174
QY 181 GCGGCTCGAGGTGCGGCTGATGTGACGACAT 214
Db 175 GCGGTGTGAGGTGCGGCTGACAGACGACAT 211

RESULT 10

AAS05211 ID AAS05211 standard; DNA; 208 BP.

AAS05211; AC

DT 07-SEP-2001 (first entry)

DE Mycobacterium marinum rpoB gene fragment.

KW Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium marinum.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism

PS Claim 1; Page 43; 50pp; English.

The present sequence for Mycobacterium marinum rpoB gene fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from various Mycobacterium species. These rpoB gene fragments can be used in the diagnosis and identification of Mycobacterium species using a novel PCR-restriction fragment length polymorphism analysis (PRA) method. The method comprises obtaining a restriction fragment length polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA fragment from the microorganism to be identified and comparing the RFLP patterns from the known rpoB gene fragments with the unidentified fragment. The rpoB gene fragments are useful to identify a wide range of Mycobacterium species, e.g. for diagnosis or to obtain epidemiological and pathogenesis information for selection of appropriate therapies, including M. tuberculosis, M. leprae and non-tuberculous mycobacteria (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and cost effective (only 1 PCR

CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

SO Sequence 208 BP; 45 A; 69 C; 68 G; 26 T; 0 other;

Query Match 65.8%; Score 140.8; DB 22; Length 208;
Best Local Similarity 82.2%; Pred. No. 4.8e-22;

Matches 176: Conservative 0; Mismatches 32; Indels 6; Gaps 1;

1 TCAGAGAGAGCGCTACGATCTGGCCCGCTGGCTGTACAGGTGAACAAGAGCTGG 60
1 TCAGAGAGAGCGCTACGATCTGGCCCGCTGGCTGTACAGGTGAACAAGAGCTGG 60
61 GCCTGGCGGCGACCAATCCGCTCAGTGCAGCACCACCCCTCAGCAGAGAGAGCTCG 120
61 GCCTGGCGGCGGCGACCAATCCGCTCAGTGCAGCACCACCCCTCAGCAGAGAGAGCTCG 114
121 TGGCCACCATGAGTACTGTGGCGCTGCACAGAGGCGCAGACCCAGATGACCGCCCG 180
115 TGGCCACCATGAGTACTGTGGCGCTGCACAGAGGCGCAGACCCAGATGACCGCTCGG 174
181 GCGGCGTGCAGGTGCGGCTGGATGTGGAGCAGCAT 214
175 GCGGCGTGCAGGTGCGGCTGGATGTGGAGCAGCAT 208

RESULT 11

AAS05217
ID AAS05217 standard; DNA; 208 BP.

AC AAS05217;

DT 07-SEP-2001 (first entry)

DE Mycobacterium celatum rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium celatum.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
XX
PS Claim 1; Page 45; 50pp; English.

CC The present sequence for Mycobacterium celatum rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments

CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

SO Sequence 208 BP; 48 A; 71 C; 64 G; 25 T; 0 other;

Query Match 65.8%; Score 140.8; DB 22; Length 208;
Best Local Similarity 82.2%; Pred. No. 4.8e-22;

Matches 176: Conservative 0; Mismatches 32; Indels 6; Gaps 1;

1 TCAGAGAGAGCGCTACGATCTGGCCCGCTGGCTGTACAGGTGAACAAGAGCTGG 60
1 TCAGAGAGAGCGCTACGATCTGGCCCGCTGGCTGTACAGGTGAACAAGAGCTGG 60
61 GCCTGGCGGCGACCAATCCGCTCAGTGCAGCACCACCCCTCAGCAGAGAGAGCTCG 120
61 GCCTGGCGGCGGCGACCAATCCGCTCAGTGCAGCACCACCCCTCAGCAGAGAGAGCTCG 114
121 TGGCCACCATGAGTACTGTGGCGCTGCACAGAGGCGCAGACCCAGATGACCGCCCG 180
115 TGGCCACCATGAGTACTGTGGCGCTGCACAGAGGCGCAGACCCAGATGACCGCTCGG 174
181 GCGGCGTGCAGGTGCGGCTGGATGTGGAGCAGCAT 214
175 GCGGCGTGCAGGTGCGGCTGGATGTGGAGCAGCAT 208

RESULT 12

AAS05202
ID AAS05202 standard; DNA; 208 BP.

AC AAS05202;

DT 07-SEP-2001 (first entry)

DE Mycobacterium gordonae type II rpoB gene fragment.

XX

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KM PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium gordonae type II.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
XX
PS Claim 1; Page 40; 50pp; English.

CC The present sequence for Mycobacterium gordonae type II rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a

DR WPI: 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
restriction fragment length polymorphism

PS Claim 1: Page 47: 50pp: English.

XX The present sequence for Mycobacterium xenopi rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC experiment), and can differentiate between many species in a single
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

SQ Sequence 208 BP: 51 A; 65 C; 67 G; 25 T; 0 other;

Query Match 63.6%; Score 136; DB 22; Length 208;
Best Local Similarity 80.8%; Pred. No. 5.3e-21;

Matches 173; Conservative 0; Mismatches 35; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGATCTGGCCGCTGGGTGCTGACAAAGTGAACAAAGACTGG 60
DB 1 TCAAGGAGAGCGCTACGATCTGGCCGCTGGGTGCTGACAAAGTGAACAAAGACTGG 60
QY 61 GCGTGGCGCGACCAATCCGGCTCAGGTGACGACCAACCCCTCAGCAGGAAGACGTGC 120
DB 61 GCGTGGCGCGACCAATCCGGCTCAGGTGACGACCAACCCCTCAGCAGGAAGACGTGC 114
QY 121 TCGCCACCATCGATGCTGCTGGCTGACAGAGGCCACGACGATGACCGCCCG 180
DB 115 TCGCCACCATCGATGCTGCTGGCTGACAGAGGCCACGACGATGACCGCCCG 174
QY 181 GCGGCGTCGAGGTGCGGCTGATGTGACGACAT 214
DB 175 GCGGCGTCGAGGTGCGGCTGATGTGACGACGACAT 208

RESULT 15

AAS05205 standard: DNA: 208 BP.

AC AAS05205;

DT 07-SEP-2001 (first entry)

DE Mycobacterium tuberculosis rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
restriction fragment length polymorphism analysis; ds.

XX Mycobacterium tuberculosis.

XX WO200131061-A1.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000MO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

XX (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
restriction fragment length polymorphism

PS Disclosure: Page 41: 50pp: English.

XX The present sequence for Mycobacterium tuberculosis rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC experiment), and can differentiate between many species in a single
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

SQ Sequence 208 BP: 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 60.6%; Score 129.6; DB 22; Length 208;
Best Local Similarity 79.0%; Pred. No. 1.3e-19;

Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGATCTGGCCGCTGGGTGCTGACAAAGTGAACAAAGACTGG 60
DB 1 TCAAGGAGAGCGCTACGATCTGGCCGCTGGGTGCTGACAAAGTGAACAAAGACTGG 60
QY 61 GCGTGGCGCGACCAATCCGGCTCAGGTGACGACCAACCCCTCAGCAGGAAGACGTGC 120
DB 61 GCGTGGCGCGACCAATCCGGCTCAGGTGACGACCAACCCCTCAGCAGGAAGACGTGC 114
QY 121 TCGCCACCATCGATGCTGCTGGCTGACAGAGGCCACGACGATGACCGCCCG 180
DB 115 TCGCCACCATCGATGCTGCTGGCTGACAGAGGCCACGACGATGACCGCCCG 174
QY 181 GCGGCGTCGAGGTGCGGCTGATGTGACGACAT 214
DB 175 GCGGCGTCGAGGTGCGGCTGATGTGACGACGACAT 208

Search completed: November 12, 2002, 16:50:25
Job time : 142.151 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 ; Search time 654.37 Seconds

(without alignments)
9517.553 Million cell updates/sec

Title: US-09-697-123b-21

Sequence: 1 tcaagagaagcgtacgacat.....ccggtgacgtgacgacat 214

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sy:*
12: gb_un:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_pl:*
25: em_ro:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pln:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vrl:*
38: em_sy:*
39: em_hlg_hum:*
40: em_hlg_mus:*
41: em_hlgc_other:*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	132.8	62.1	3752	1	MSU24494
2	129.6	60.6	610	1	MTU318818
3	129.6	60.6	610	1	MTU318819
4	129.6	60.6	616	1	MTU318812
5	129.6	60.6	618	1	MTU318815
6	129.6	60.6	618	1	MTU318817
7	129.6	60.6	633	1	MTU318814
8	129.6	60.6	637	1	MTU318816
9	129.6	60.6	639	1	MTU318821
10	129.6	60.6	970	6	AX111339
11	129.6	60.6	3534	6	AX111339
12	129.6	60.6	3853	1	MTU12205
13	129.6	60.6	5084	1	MSGP08
14	129.6	60.6	19352	1	AE006964
15	129.6	60.6	19770	1	MTCI376
16	108.8	50.8	3447	6	AR067447
17	108.8	50.8	37617	1	MTB1790C
18	108.8	50.8	34850	1	MEPRTN7
19	95.2	44.5	3941	1	AF242549
20	80.4	37.6	3495	6	AX120631
21	80.4	37.6	328050	1	AP005275
22	80.4	37.6	349980	6	AX127144
23	76	35.5	32923	1	SCD82
24	46.2	21.6	1672	1	RCPEPTR
25	46	21.5	166050	1	AL646085
26	44.6	20.8	34023	1	SC2G2
27	44.4	20.7	190050	1	AL646059
28	44.2	20.7	8611	8	SCU74495
29	44	20.6	776	5	CHKTNC
30	44	20.6	39531	1	SCP8
31	43.8	20.5	1782	14	AF171937
32	43.8	20.5	11103	1	AE001944
33	43.2	20.2	1813	8	ME2NIRE
34	43	20.1	101365	2	AP004748
35	43	20.1	152263	2	AP005412
36	42.8	20.0	29625	1	SCB22A
37	42.8	20.0	77534	1	AF235504
38	42.6	19.9	24700	1	SCB8
39	42.6	19.9	113193	1	AF357202
40	42.6	19.9	133066	2	CNS08C80
41	42.4	19.8	10660	1	AE001963
42	42.2	19.7	1665	14	AF403050
43	42.2	19.7	1674	14	AF403049
44	42.2	19.7	1734	6	104901
45	42.2	19.7	1734	6	108509

ALIGNMENTS

RESULT 1
MSU24494
LOCUS MSU24494
DEFINITION Mycobacterium smegmatis DNA polymerase (rpoB) gene, complete cds.
ACCESSION U24494
VERSION U24494.1 GI:790347
KEYWORDS
SOURCE
ORGANISM Mycobacterium smegmatis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
REFERENCE
1 (bases 1 to 3752)
Hetherington, S.V., Watson, A.S. and Patrick, C.C.
Sequence and analysis of the rpoB gene of Mycobacterium smegmatis

Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE

1 Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL

Unpublished
2 (bases 1 to 610)

REFERENCE

Unpublished

AUTHORS

Herrera, L.

TITLE

Direct Submission

JOURNAL

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriología, Centro Nacional Microbiología, Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid. 28220, SPAIN

FEATURES

source
1. 610
Location/Qualifiers

gene
/organism="Mycobacterium tuberculosis"
/isolate="1417-97"
/db_xref="taxon:1773"
1. 610
/gene="rpoB"
<1..>610
/codon_start=1
/transl_table=1
/product="RNA polymerase beta subunit"
/protein_id="CAC87036.1"
/db_xref="GI:22208415"
/translation="LDYRKLRGEPPTKESAOITLLENLFKEKRYDLARVGRYKVK
KLGLHVEPTITSTLTEDVAVATLEYLRLHEQOTMTVPGVEVEVDIDHFGNR
RLRTVGEIIONIRVGMSEMRVREMTQDVEATPOTLIIRPVVAIKKEFGTS
QPSQFMQNNPLSGLTGHRRLSALGPGLSRERAGLEVDRVHP"

CDS

BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN

Query Match

Best Local Similarity 60.68; Score 129.6; DB 1; Length 610;
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY 1 TCAGAGAGAGCGCTACATCTGCGCCGCGTGCCTGACAGAGTGAACAGAGCTGG 60
DB 80 TCAGAGAGAGCGCTACATCTGCGCCGCGTGCCTGACAGAGTGAACAGAGCTGG 60
QY 61 GCGTGGCGGACCAATCCGCTCAGGTGACACACACCTTCACCGAGAGAGCTGG 120
DB 140 GCGTGGCGGACCAATCCGCTCAGGTGACACACACCTTCACCGAGAGAGCTGG 120
QY 121 TCGCCACCATGAGTACCTGCTGCGCTGCACAGAGGCGACACAGATGACCGCCCGG 180
DB 194 TCGCCACCATGAGTACCTGCTGCGCTGCACAGAGGCGACACAGATGACCGCCCGG 180
QY 181 CCGGCGTGCAGAGTCCGCTGATGTGAGCAGACAT 214
DB 254 GCGGCGTGCAGAGTCCGCTGATGTGAGCAGACAT 214
QY 254 GCGGCGTGCAGAGTCCGCTGATGTGAGCAGACAT 214
DB 254 GCGGCGTGCAGAGTCCGCTGATGTGAGCAGACAT 214

RESULT 4

MTU318813

LOCUS

616 bp DNA

DEFINITION

Mycobacterium tuberculosis partial rpoB gene for RNA polymerase beta subunit, isolate 1763-97.

ACCESSION

AJ318813

VERSION

1

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

Mycobacterium tuberculosis

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE

1 Herrera, L., Jimenez, M.S. and Saez, J.A.

AUTHORS

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB

TITLE

Unpublished

JOURNAL

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriología, Centro

gene and review

REFERENCE

Unpublished

AUTHORS

Herrera, L.

TITLE

Direct Submission

JOURNAL

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriología, Centro Nacional Microbiología, Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid. 28220, SPAIN

FEATURES

source
1. 616
Location/Qualifiers

gene
/organism="Mycobacterium tuberculosis"
/isolate="1763-97"
/db_xref="taxon:1773"
1. 615
/gene="rpoB"
<1..>615
/codon_start=1
/transl_table=1
/product="RNA polymerase beta subunit"
/protein_id="CAC87030.1"
/db_xref="GI:22208403"
/translation="LDYRKLRGEPPTKESAOITLLENLFKEKRYDLARVGRYKVK
KLGLHVEPTITSTLTEDVAVATLEYLRLHEQOTMTVPGVEVEVDIDHFGNR
RLRTVGEIIONIRVGMSEMRVREMTQDVEATPOTLIIRPVVAIKKEFGTS
QPSQFMQNNPLSGLTGHRRLSALGPGLSRERAGLEVDRVHP"

CDS

BASE COUNT 125 a 191 c 201 g 99 t
ORIGIN

Query Match

Best Local Similarity 60.68; Score 129.6; DB 1; Length 616;
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY 1 TCAGAGAGAGCGCTACATCTGCGCCGCGTGCCTGACAGAGTGAACAGAGCTGG 60
DB 80 TCAGAGAGAGCGCTACATCTGCGCCGCGTGCCTGACAGAGTGAACAGAGCTGG 60
QY 61 GCGTGGCGGACCAATCCGCTCAGGTGACACACACCTTCACCGAGAGAGCTGG 120
DB 140 GCGTGGCGGACCAATCCGCTCAGGTGACACACACCTTCACCGAGAGAGCTGG 120
QY 121 TCGCCACCATGAGTACCTGCTGCGCTGCACAGAGGCGACACAGATGACCGCCCGG 180
DB 194 TCGCCACCATGAGTACCTGCTGCGCTGCACAGAGGCGACACAGATGACCGCCCGG 180
QY 181 CCGGCGTGCAGAGTCCGCTGATGTGAGCAGACAT 214
DB 254 GCGGCGTGCAGAGTCCGCTGATGTGAGCAGACAT 214
QY 254 GCGGCGTGCAGAGTCCGCTGATGTGAGCAGACAT 214
DB 254 GCGGCGTGCAGAGTCCGCTGATGTGAGCAGACAT 214

RESULT 5

MTU318815

LOCUS

618 bp DNA

DEFINITION

Mycobacterium tuberculosis partial rpoB gene for RNA polymerase beta subunit, isolate 2540-97.

ACCESSION

AJ318815

VERSION

1

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

Mycobacterium tuberculosis

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE

1 Herrera, L., Jimenez, M.S. and Saez, J.A.

AUTHORS

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB

TITLE

Unpublished

JOURNAL

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriología, Centro

Unpublished

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriología, Centro

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriología, Centro

Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN

FEATURES

Location/Qualifiers

source

1..618
/organism="Mycobacterium tuberculosis"
/isolate="2540197"
/db_xref="taxon:1773"

gene

1..618
/gene="rpoB"

CDS

<1..>618
/gene="rpoB"

/codon_start=1
/transl_table=11

/product="RNA polymerase beta subunit"

/protein_id="CAC87032.1"

/db_xref="GI:22208407"

/translation="TDEALDLYRKLRLRGPPTKESQTLLENLFEKRYDLARYGR
YVNNKRLHVEPTTSSLTLEEDVVAITEYLVLHGGTTMTVPGVEVPEVETDID
HFGNRLRTVGLIONOIRVGMSSRMERYRREMTODVEATTPRTLINRPVVAIKKE
FPGTSOLSGFMQNNPLSLGLTHKRRLLALPGSLSRERAGLEVADV"

BASE COUNT

124 a 190 c 207 g 97 t

ORIGIN

Query Match

Best Local Similarity 79.0%; Score 129.6; DB 1; Length 618;

Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY

1 TCAAGAGAAGCGCTACGATCTGGCCGCGGTGCTGCTACAGGTGAACAAGACTCG 60
|||||

Db

95 TCAAGAGAAGCGCTACGACTGGCCGCGGTGCTGCTATTAAGTCAACAAGAGCTCG 154
|||||

QY

61 GCGTGGCGGCAACATCCGGCTCAGGTGACCAACCAACCCCTGACGAGAGAGCTCG 120
|||||

Db

155 GCGTGCATGTCCGCGAGCC-----CATCACGTCGTCGACGCTGACGAGAAGAGCTCG 208
|||||

QY

121 TCGCCACATCGAGTACCTGCTGCGCTGACAGAGGCGGACAGCAAGATGACCCCGG 180
|||||

Db

209 TGGCCACATCGAATATCTGGTCCCTTGACAGAGGCTGACAGCATGACGTTCCGG 268
|||||

QY

181 GCGGCGTCGAGGTGCGCGGTGATGTGACGACAT 214
|||||

Db

269 GCGGCGTCGAGGTGCGCGGTGAAACGACGACAT 302
|||||

RESULT 6

MTU318817 618 bp DNA circular BCT 09-AUG-2002

LOCUS

MycoBacterium tuberculosis partial rpoB gene for RNA polymerase

DEFINITION

beta subunit, isolate 2348-98.

ACCESSION

AJ318817

VERSION

GI:22208410

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

MycoBacterium tuberculosis

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE

1

AUTHORS

Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 618)

AUTHORS

Herrera, L.

TITLE

Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN

JOURNAL

Location/Qualifiers

FEATURES

1..618
/organism="Mycobacterium tuberculosis"
/isolate="2348-98"

/db_xref="taxon:1773"
1..618
/gene="rpoB"

CDS

<1..>618
/gene="rpoB"

/codon_start=1
/transl_table=11

/product="RNA polymerase beta subunit"

/protein_id="CAC87034.1"

/db_xref="GI:22208411"

/translation="TDEALDLYRKLRLRGPPTKESQTLLENLFEKRYDLARYGR
YVNNKRLHVEPTTSSLTLEEDVVAITEYLVLHGGTTMTVPGVEVPEVETDID
HFGNRLRTVGLIONOIRVGMSSRMERYRREMTODVEATTPRTLINRPVVAIKKE
FPGTSOPSPQFMQNNPLSLGLTHKRRLLALPGSLSRERAGLEVADV"

BASE COUNT

124 a 192 c 207 g 95 t

ORIGIN

Query Match

Best Local Similarity 79.0%; Score 129.6; DB 1; Length 618;

Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY

1 TCAAGAGAAGCGCTACGATCTGGCCGCGGTGCTGCTACAGGTGAACAAGACTCG 60
|||||

Db

95 TCAAGAGAAGCGCTACGACTGGCCGCGGTGCTGCTATTAAGTCAACAAGAGCTCG 154
|||||

QY

61 GCGTGGCGGCAACATCCGGCTCAGGTGACCAACCAACCCCTGACGAGAGAGCTCG 120
|||||

Db

155 GCGTGCATGTCCGCGAGCC-----CATCACGTCGTCGACGCTGACGAGAAGAGCTCG 208
|||||

QY

121 TCGCCACATCGAGTACCTGCTGCGCTGACAGAGGCGGACAGCAAGATGACCCCGG 180
|||||

Db

209 TGGCCACATCGAATATCTGGTCCCTTGACAGAGGCTGACAGCATGACGTTCCGG 268
|||||

QY

181 GCGGCGTCGAGGTGCGCGGTGATGTGACGACAT 214
|||||

Db

269 GCGGCGTCGAGGTGCGCGGTGAAACGACGACAT 302
|||||

RESULT 7

MTU318814 633 bp DNA circular BCT 09-AUG-2002

LOCUS

MycoBacterium tuberculosis partial rpoB gene for RNA polymerase

DEFINITION

beta subunit, isolate 1058-97.

ACCESSION

AJ318814

VERSION

GI:22208404

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

MycoBacterium tuberculosis

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE

1

AUTHORS

Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 633)

AUTHORS

Herrera, L.

TITLE

Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN

JOURNAL

Location/Qualifiers

FEATURES

1..633

source

/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"

gene

1..633
/gene="rpoB"

CDS

<1..>633
/gene="rpoB"

/codon_start=1

```

                                /transl_table=11
                                /product="RNA polymerase beta subunit"
                                /protein_id="CAC87031.1"
                                /db_xref="GI:22208405"
                                /translation="DNIVGTDEALLDIYRKLRPGESPPTKESAQTLLLENLFFKREKRYDL
                                AVNGRKVNKKLGLHVGEPITSSLTLEEDVATIEYLRLHGGOTTMTVPGCVPEVPE
                                TDDIHFGNRRRLRTVGEILIONIRVGMSEMERVVERMTTODVEAITPOTLIINRPVY
                                AAIKEFFGTSQISQFMDONNPISGLTYKRRISALGPGLSRERAGLEVRDY"
BASE COUNT      129 a      195 c      210 g      99 t
ORIGIN
Query Match      60.6%; Score 129.6; DB 1; Length 633;
Best Local Similarity 79.0%; Pred. No. 1.6e-12;
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

OY 1 TCAAGAGAAAGCGCTACGATCTGGCGCCGCGGTGCTACAGGTGAACGAAGAAGCTGG 60
    |||||||
DB 110 TCAAGAGAAAGCGCTACGACCTGGCCCGCGGTGCTATTAAGTCAACAAGAAGCTCG 169
    |||||||

OY 61 GCCTGGCGGCAACCAATCCGCTCAAGTGACCAACACCTCAACCGAAGAAAGCTGG 120
    |||||
DB 170 GCGTGCATGTGCGGCGAGCC-----CATCAGCTGCTGCAACGCTGACCGAAGAAGACGCTG 223
    |||||

OY 121 TCGCCACCATCGAGTACCTGGTGGCGCTGCACGAGGGCCAGACACGATGACCGCCCGG 180
    |||||||
DB 224 TGGCCACCATCGAATATCTGCTGCTGCTGACAGAGGTCAGACACGATGACCGTCCGG 283
    |||||||

OY 181 GCGGCGTGCAGGTGCGGCTGATGTGAGACGACAT 214
    |||||||
DB 284 GCGGCGTGCAGGTGCGGCTGGAACCGACGACAT 317
    |||||||

RESULT 8
MTU318816      637 bp      DNA      circular BCT 09-AUG-2002
LOCUS
DEFINITION
beta subunit, isolate 1253-98.
ACCESSION
AJ318816
VERSION
AJ318816.1 GI:22208408
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera, L., Jimenez, M. S. and Saez, J. A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 637)
AUTHORS
Herrera, L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
National Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN
FEATURES
source
1. .637
/organism="Mycobacterium tuberculosis"
/isolate="1253-98"
/db_xref="taxon:1773"
1. .637
/gene="rpoB"
<1. .>637
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87033.1"
/db_xref="GI:22208409"
/translation="VGTDEALLDIYRKLRPGESPPTKESAQTLLLENLFFKREKRYDLARY
GRAYKNKKLGLHVGEPITSSLTLEEDVATIEYLRLHGGOTTMTVPGCVPEVETDD
IHFGRNRRRLRTVGEILIONIRVGMSEMERVVERMTTODVEAITPOTLIINRPVAAI
KEFFGTSQISQFMDONNPISGLTYKRRISALGPGLSRERAGLEVRDYHPSH"
BASE COUNT      126 a      202 c      212 g      99 t
ORIGIN
Query Match      60.6%; Score 129.6; DB 1; Length 639;
```

```

                                IDHFGNRRRLRTVGEILIONIRVGMSEMERVVERMTTODVEAITPOTLIINRPVAAI
                                KEFFGTSQISQFMDONNPISGLTYKRRISALGPGLSRERAGLEVRDYHPSH"
BASE COUNT      128 a      198 c      210 g      101 t
ORIGIN
Query Match      60.6%; Score 129.6; DB 1; Length 637;
Best Local Similarity 79.0%; Pred. No. 1.6e-12;
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

OY 1 TCAAGAGAAAGCGCTACGATCTGGCGCCGCGGTGCTACAGGTGAACGAAGAAGCTGG 60
    |||||||
DB 101 TCAAGAGAAAGCGCTACGACCTGGCCCGCGGTGCTATTAAGTCAACAAGAAGCTCG 160
    |||||||

OY 61 GCCTGGCGGCAACCAATCCGCTCAAGTGACCAACACCTCAACCGAAGAAAGCTGG 120
    |||||
DB 161 GCGTGCATGTGCGGCGAGCC-----CATCAGCTGCTGCAACGCTGACCGAAGAAGACGCTG 214
    |||||||

OY 121 TCGCCACCATCGAGTACCTGGTGGCGCTGCACGAGGGCCAGACCATGACCGCCCGG 180
    |||||||
DB 215 TGGCCACCATCGAATATCTGCTGCTGCTGACAGAGGTCAGACACGATGACCGTCCGG 274
    |||||||

OY 181 GCGGCGTGCAGGTGCGGCTGATGTGAGACGACAT 214
    |||||||
DB 275 GCGGCGTGCAGGTGCGGCTGGAACCGACGACAT 308
    |||||||

RESULT 9
MTU318821      639 bp      DNA      circular BCT 09-AUG-2002
LOCUS
DEFINITION
beta subunit, isolate 1071-98.
ACCESSION
AJ318821
VERSION
AJ318821.1 GI:22208418
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
AUTHORS
Herrera, L., Jimenez, M. S. and Saez, J. A.
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 639)
AUTHORS
Herrera, L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
National Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN
FEATURES
source
1. .639
/organism="Mycobacterium tuberculosis"
/isolate="1071-98"
/db_xref="taxon:1773"
1. .639
/gene="rpoB"
<1. .>639
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87038.1"
/db_xref="GI:22208419"
/translation="RRTDEALLDIYRKLRPGESPPTKESAQTLLLENLFFKREKRYDLARY
GRAYKNKKLGLHVGEPITSSLTLEEDVATIEYLRLHGGOTTMTVPGCVPEVETDD
IDHFGNRRRLRTVGEILIONIRVGMSEMERVVERMTTODVEAITPOTLIINRPVAAI
KEFFGTSQISQFMDONNPISGLTYKRRISALGPGLSRERAGLEVRDYHPSH"
BASE COUNT      126 a      202 c      212 g      99 t
ORIGIN
Query Match      60.6%; Score 129.6; DB 1; Length 639;
```


Best Local Similarity 79.0%; Pred. No. 1.6e-12;
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY 1 TCAAGAGAACGCTACGATCTGCCCCGCGTGGCTGCTACACGTAACAGAACCTCG 60
|||||
Db 101 TCAAGAGAACGCTACGATCTGCCCCGCGTGGCTGCTATAGGTCAACAGAACCTCG 160
61 GCTGGGGGGGACCAATCCGGCTCAGGTGACCAACACCACTCCACAGAGAACGCTCG 120
|||||
Db 161 GCTGTCATGTGCGGACGCC-----CATCACGTGCTGACAGGTCACCAAGAACGCTCG 214
QY 121 TCGCCACCATCAGTACCTGCTGCGCTGCACGAGGCGCACGACGATGACCGCCCG 180
|||||
Db 215 TGGCCACCATCAGTATCTGCTGCGCTTGCACGAGGCTCAGACGATGACCGCTTCGG 274
QY 181 GCGGCGTCGAGCTCCGCTGATGTGACGACAT 214
|||||
Db 275 GCGGCGTCGAGGTCCGCTGGAACCGACGACAT 308

RESULT 10
150706 150706 970 bp DNA linear PAT 07-OCT-1997
LOCUS Sequence 1 from patent US 5643723.
ACCESSION 150706
VERSION 150706.1 GI:2472409
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 970)
AUTHORS Persing, D.H., Hunt, J.J., Young, K.K.Y., Felmlie, T.A., Roberts, G.D.
TITLE Detection of a genetic locus encoding resistance to rifampin in
mycobacterial cultures and in clinical specimens
JOURNAL Patent: US 5643723-A 1 01-JUL-1997;
FEATURES
source 1..970
BASE COUNT 182 a 302 c 330 g 156 t
ORIGIN

Query Match 60.6%; Score 129.6; DB 6; Length 970;
Best Local Similarity 79.0%; Pred. No. 1.5e-12;
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY 1 TCAAGAGAACGCTACGATCTGCCCCGCGTGGCTGCTACACGTAACAGAACCTCG 60
|||||
Db 26 TCAAGAGAACGCTACGATCTGCCCCGCGTGGCTGCTATAGGTCAACAGAACCTCG 85
QY 61 GCTGGGGGGGACCAATCCGGCTCAGGTGACCAACACCACTCCACAGAGAACGCTCG 120
|||||
Db 86 GCTGTCATGTGCGGACGCC-----CATCACGTGCTGACAGGTCACCAAGAACGCTCG 139
QY 121 TCGCCACCATCAGTACCTGCTGCGCTGCACGAGGCGCACGACGATGACCGCCCG 180
|||||
Db 140 TGGCCACCATCAGTATCTGCTGCGCTTGCACGAGGCTCAGACGATGACCGCTTCGG 199
QY 181 GCGGCGTCGAGGTCCGCTGATGTGACGACAT 214
|||||
Db 200 GCGGCGTCGAGGTCCGCTGGAACCGACGACAT 233

RESULT 11
AX11339 3534 bp DNA linear PAT 30-APR-2001
LOCUS Sequence 2072 from Patent WO0123604.
ACCESSION AX11339
VERSION AX11339.1 GI:13927631
KEYWORDS
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE 1 (bases 1 to 3534)
AUTHORS Bergeron, M.G., Boissinot, M., Huiletsky, A., m Nard, C., Ouellette, M.,
Picard, F.J., and Roy, P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 2072 05-APR-2001;
FEATURES
source Infectio Diagnostic (I.D.I.) INC. (CA)
Location/Qualifiers
1..3534
/organism="Mycobacterium tuberculosis"
/strain="Rv"
/db_xref="taxon:1773"
BASE COUNT 679 a 1081 c 1188 g 586 t
ORIGIN

Query Match 60.6%; Score 129.6; DB 6; Length 3534;
Best Local Similarity 79.0%; Pred. No. 1.3e-12;
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY 1 TCAAGAGAGAGCGCTACGATCTGCCCCGCGTGGCTGCTACACGTAACAGAACCTCG 60
|||||
Db 902 TCAAGAGAGAGCGCTACGATCTGCCCCGCGTGGCTGCTATAGGTCAACAGAACCTCG 961
QY 61 GCTGGGGGGGACCAATCCGGCTCAGGTGACCAACCACTCCACAGAGAACGCTCG 120
|||||
Db 962 GCTGTCATGTGCGGACGCC-----CATCACGTGCTGACAGGTCACCAAGAACGCTCG 1015
QY 121 TCGCCACCATCAGTACCTGCTGCGCTGCACGAGGCGCACGACGATGACCGCCCG 180
|||||
Db 1016 TGGCCACCATCAGTATCTGCTGCGCTTGCACGAGGTCACGACGATGACCGCTTCGG 1075
QY 181 GCGGCGTCGAGGTCCGCTGATGTGACGACAT 214
|||||
Db 1076 GCGGCGTCGAGGTCCGCTGGAACCGACGACAT 1109

RESULT 12
MTU12205 3853 bp DNA linear BCT 02-MAR-2000
LOCUS Mycobacterium tuberculosis H37Rv RNA-polymerase beta subunit (rpoB)
DEFINITION gene, partial cds.
ACCESSION U12205
VERSION U12205.1 GI:515684
KEYWORDS
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Imboden, P., Troller, R., Marchesi, F., Telenti, A., Bodmer, T.,
Cole, S., Schopfer, K., and Burkart, T.
TITLE The rpoB gene of Mycobacterium tuberculosis
JOURNAL Unpublished
AUTHORS Imboden, P.
REFERENCE 2 (bases 1 to 3853)
TITLE Direct Submission
SUBMITTED (11-JUL-1994) Paul Imboden, Institute for Medical
Microbiology, University of Berne, Friedbuehlstrasse 51, Berne,
3010, Switzerland
FEATURES
source Location/Qualifiers
1..3853
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
576..>3853
/gene="rpoB"
576..>3853
/gene="rpoB"
/codon_start=1
/transl_table=11

[illegible]

gene
COS

AFWREVNRGALDTEGPTTSGATCADNRHVLLTVDDAGHRMPSFATOTLWRFPA
AHRP*
11859..13487
/gene="MT0701"
11859..13487
/note="similar to SP:P33224 GB:L20915 PID:457172
PID:457174 PID:537028; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="acyl-CoA dehydrogenase, putative"
/protein_id="AAK44926.1"
/db_xref="GI:13880223"
/translation="MSDTHVYTNVNPPLNYPASSPYLIPALIOEGCGNGLDEVN
GASASCOBMRGELADNRNPLTHHDAVGRVDEVEDPAYHELMKTAITHGHAAP
MADRRGAVHVAKTSYVTERPHICISMTYAVPALRINSELAAYEPPLTSREY
DDELKATTKKATITGMSMTERKGGSDVRACTQATPADGSYSLTGHKFTSAPMD
ITVLQAAPDGLSCFLPRVLPDGTNRNMLQRLDKIGNHANSSEVEYDAVAMLY
GEGRGVPTIIEVNLTRLDALGATSMTGLRAVHAQHRKAFGAYLIDPLMRN
VLADLAVEAEATIVAMRAGATONAVRGNETEALRRIGLAARKYVWCKRSTAAAE
ALECLGNGYVEDSGMPRIYREAPLMGIWEGSGVSLDITRAMATRPACVEVLFDEL
ARAGODPRLDGVRRLRPPOLDITGYRARKIAEDICIALGSLLVRRGHFAVAAE
FLATRLGGONGGAVTNPAGLDLAPILERALVKG"

gene
CDS

13498..14436
/gene="MT0702"
13498..14436
/gene="MT0702"
/note="similar to GP:3885480; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="enoyl-CoA hydratase/isomerase family protein"
/protein_id="AAK44927.1"
/db_xref="GI:13880224"
/translation="MTHAIRVDFDNLKMTMYEYGRHARTFNRPEKNAITADPL
ELSLVERADLDPCGVHILVSGRGEFCAGFDLSAYHEGSSVGGGAGTGLDCKT
QAVNHLPWOPMDIDYQMSRFRGASLHMDKPTVATHEYVAGATGCTVLDLKT
VIAAADKIGTPTRVWGVPAAGLMAHRLGDORAKRLFLFGDCTTGAQAEMGLAVEA
PEPADLKERTELVAIALPVNQLIMVKLALNSALLOQCVATSRMVSFTDCAARHT
PEGHAFVADAVEHGFDAVRRDPEFDYGRQASRV"
14439..15161
/gene="MT0703"
14439..15161
/gene="MT0703"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"

Query Match 60.6%; Score 129.6; DB 1; Length 19352;
Best Local Similarity 79.0%; Pred. No. 1e-12;
Matches 169; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

QY 1 TCAGAGAGAGCGCTAGCATCTGCGCCGCTGCTGCTACAGCAAGCAAGCTGG 60
|||||
DB 1064 TCAGAGAGAGCGCTAGCATCTGCGCCGCTGCTGCTACAGCAAGCAAGCTGG 1123
|||||
QY 61 GCTGGGGGGCAACCAATCGGCTCAGGTGACACACCTCACCAGGAAGAGCTCG 120
|||||
DB 1124 GCTGGGGGGCAACCAATCGGCTCAGGTGACACACCTCACCAGGAAGAGCTCG 1177
|||||
QY 121 TCGCCACCATGAGTACCTGCTGCGCCGCGGCGGCGACGATACAGTACCGCCCGG 180
|||||
DB 1178 TCGCCACCATGAGTACCTGCTGCGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1237
|||||
QY 181 GCGGCGTGAAGTCCGCTGAGTGTGAGACAT 214
|||||
DB 1238 GCGGCGTGAAGTCCGCTGAGTGTGAGACAT 1271
|||||

RESULT 15
MTC1376

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

MTC1376 19770 bp DNA linear BCT 03-AUG-2001
Mycobacterium tuberculosis H37Rv complete genome; segment 32/162.
295972.1 AL123456
295972.1 GI:3261190

REFERENCE
AUTHORS

Mycobacterium tuberculosis H37Rv.
Mycobacterium tuberculosis H37Rv.
Actinomycetales: Corynebacteriaceae; Mycobacteriaceae;
Mycobacterium tuberculosis complex.
1 (bases 1 to 19770)
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.,
Harris D., Gordon S.V., Eigmeier K., Gas S., Barry III C.E.,
Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
Connor R., Davies R., Devlin K., Fellwell T., Gentles S.,
Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J.,
Moule S., Murphy L., Oliver S., Osborne J., Quail M.A.,
Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S.,
Squires S., Squires R., Sulston J.E., Taylor K., Whitehead S. and
Barrell B.G.

TITLE

Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence

JOURNAL
MEDLINE
PUBMED

Nature 393 (6685), 537-544 (1998)
98295987
9634230

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

2 (bases 1 to 19770)
Parkhill J.
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unité de Génétique
Moléculaire Bactérienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 27, 1998 this sequence version replaced gi:2143285.
Notes:

Details of M. tuberculosis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL: <http://www.sanger.ac.uk/projects/M.tuberculosis/>) CDS have
been renumbered from the original cosmid submissions but the old
gene designations are in brackets after the new gene numbers.
Gene prediction was based on a Hidden Markov Model of TB genes.
Implemented in TParse (Krogh) supplemented with visual inspection
of positional base preference in codons, especially where there is
an increase in the observed/expected third position G + C.
CAUTION: In some cases we may not have predicted the correct
initiation codon. Where possible we choose an initiation codon
(atg, gtg, or ttg) which is preceded by an upstream ribosome
binding site sequence (optimally 5-13bp before the initiation
codon). If this cannot be identified we choose the most upstream
initiation codon.

FEATURES
source

Location/Qualifiers

1..19770
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
/clone="1376"
57..61
/note="agga, possible rbs upstream of Rv0654"
68..1573
/gene="Rv0654"
68..1573
/note="Rv0654"
/note="O53353 LIGNOSTILBERE-ALPHA, BETA-DIOXGENASE (485
aa) opt:280 z-score:330.1 E(1):2.3e-11, (28.5% identity
in 523 aa overlap). Also similar to M. tuberculosis protein
MTCy21C12.07c (29.5% identity in 522 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0654"
/protein_id="CA809380.1"
/db_xref="GI:2143307"
/db_xref="SPTREMBL:O06785"

```
/translation="MTTQAESQNPYLEGFLAPVSTEVATDLPVGTGRIPHLDCRY
LRNGPRPAVEVPDPTATHTFTGDAMVHGVALRKGKAWIKNNRVRTPAVCAALGEPISA
RHPRTGIIEGGNTNTVLTHAGRTTLAVAGVANNELTDELDTVPCDFDGLHGGYT
AHPQDPHTGELHAVSYSFARGHVQVSVITGDGHARTRDIEVAGSPMHFSFLTDN
VYVIDLPVTFPDMQVSPASVRMLQPARLYIQSVIGVRIIPDIIALGNMOCHSD
RLPYANPSPYPARVGVMPREGNEDYRMPDIEPCVYHPLNAYSECRRACAEVYIVDYV
RYSRMEDRRGRGDSRPSLDRTMTNLATGATVTECRDRAOERPRINETIYVGPPIR
FAYTNGIEGFLVAGAAALSTPLYKODCTGSGTSTASLDPDLLIGEMFVFPVPSRAE
DGGIIMGTGMRHKGREGQLLLDAQTLSTIAIVHLPORVPMGFHGNMAPTT"
1585. .2664
/gene="Rv0655s"
/gene="Rv0655s"
1585. .2664
/note="Rv0655s"
/pasta score: YRBF_ECOLI P45393 hypothetical abc
transporter acp-binding (269 aa) opt: 644 z-score: 721.8
E(): 3.4e-33 (38.5% identity in 244 aa overlap); contains
PS00017 ATP/GTP-binding site motif A, PS00211 ABC
transporters family signature, highly similar to M. leprae
MKL MYCLE P30769 possible ribonucleotide transport atp
(347 aa) opt: 2021 z-score:2244.4 E(): 0, (92.2% identity
in 335 aa overlap). Also similar to many otherM.
tuberculosis ABC transporters eg. MTCY253.24 (33.6%
identity in 241 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0655s"
/protein_id="CAB09379.1"
/db_xref="GI:2143306"
/db_xref="SWISS-PROT:006784"
/translation="MRYSDSYHTGRMQPRASTEGFPMGVSIEVNGLTRKSPSSRIWE
DVTLLPAGAEVSVLLGPGSTGKSVFLSLGILRBERGSIILIDGDIIECSAKELYEI
RFLFGVLFDQDGLFGSMNYDNTAPLREHRTKRKSEIRDIYMEKLAIVGIGDEKRF
PGEISGMRKAGRLARALYLDQIILCEPDSGLDVPRTATLSQILIMQIDDTIL
IYTHININARTYVDNMGMLFRKHLVMEGPREVLTSDPVRQFLNGRRIGIPIGSEE
KDEATMAEQALLDGHAGVEIEGVPQISATPMPERKAVARQARVEMLHTL
PKKQAAIILDLDEGTHKAYVHEIGQ"
1762. .1785
/gene="Rv0655s"
/note="PS00017 ATP/GTP-binding site motif A"
2074. .2118
/gene="Rv0655s"
/note="PS00211 ABC transporters family signature"
complement(3052. .3435)
/gene="Rv0656c"
complement(3052. .3435)
/gene="Rv0656c"
/note="Rv0656c, (MTCI376.20, unknown), len: 127 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0656c"
/protein_id="CAB09378.1"
/db_xref="GI:2143305"
/db_xref="SPTREMBL:006783"
/translation="MAAATYTTGTHRGLELRAAQRVSCPEPORAFCRSARNADFPDQ
MSRMFGDVPVPPVPPKSVYRWIDSQHRRLARAGAVGALVVDLLICTAARGLVLH
DDADYELAEHRLPDI RVRVRSVADD"
complement(3530. .3685)
/gene="Rv0657c"
complement(3530. .3685)
/gene="Rv0657c"
complement(3530. .3685)
/note="Rv0657c, (MTCI376.19), unknown, len: 51 aa; similar
to several other M. tuberculosis hypothetical proteins eg.
YW08_MYCTU Q10848 hypothetical 8.9 kd protein cy39.08c (80
aa), fasta scores: opt: 107 z-score: 187.3 E(): 0.0038,
45.8% identity in 48 aa overlap. Also similar to MTCY48.5
andAL020958ISC4H8.7 Streptomyces coelicolor cosmid 4H8 (56
aa), 41.0% identity in 39 aa overlap."
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0657c"
/protein_id="CAB09377.1"
/db_xref="GI:2143304"
```

```
/db_xref="SPTREMBL:006782"
/translation="MSVTQIDLDDEALADVRIAAVHTKKEAVNLAMRDYVERFRRIE
ALASRP"
complement(3691. .3696)
/note="possible RBS upstream of Rv0657c"
complement(3761. .4477)
/gene="Rv0658c"
complement(3761. .4477)
/gene="Rv0658c"
/note="Rv0658c, (MTCI376.18), len: 238, unknown,
probablemembrane protein, similar to YPRB_ECOLI P33774
hypothetical124.3 kd protein (urf 1) (217 aa), fasta
scores: opt: 174 z-score: 215.8 E(): 5.3e-05, (25.6%
identity in 223 aa overlap). Also similar to MTCY359.10
(28.7% identity in 178 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0658c"
/protein_id="CAB09376.1"
/db_xref="GI:2143303"
/db_xref="SPTREMBL:006781"
/translation="MEAGRADTVAPSHRMGLGAFIVELFVASTSLAVLTGHCYV
SAGVALALAPRVYVAGLAIIILTRLRGNLRDRLRMSRKRLRLGMRFGGLMVT
IPASVYTAIVGPEANSAAVRIIFGVNASPMALVFLVVFVAPLCEETIYGLLHG
AVDRMRMALVTVTVFALHLEFARAPLVVAIPALARFSGGLASIVTHQV
TNLLPGIVLLGLGALISLP"
complement(4480. .4483)
/note="possible RBS upstream of Rv0658c"
complement(4753. .5061)
/gene="Rv0659c"
complement(4753. .5061)
/gene="Rv0659c"
complement(4753. .5061)
/note="Rv0659c, (MTCI376.17), len: 102; unknown, similar
to YW28_MYCTU Q10867 hypothetical 12.3 kd protein cy39.28
(114 aa), fasta scores: opt: 144 z-score: 213.2 E():
7.3e-05, 30.8% identity in 107 aa overlap. Also similar to
MTCY09F9.22 (32.7% identity in 101 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0659c"
/protein_id="CAB09387.1"
/db_xref="GI:2143302"
/db_xref="SPTREMBL:006780"
/translation="MRGELMFAATPGGDRPVLYLTRDYADRIGAVVVALTRRRG
LVSELELTAVERNVPSCVVFNDNIHTLPTAFRRRITRLSPARLHEACQTLRASTGC
"
complement(5048. .5293)
/gene="Rv0650c"
complement(5048. .5293)
/gene="Rv0650c"
/note="Rv0660c, (MTCI376.16), len: 81, some similarity to
IAF016485_130 Halobacterium sp: NRC-1 plasm (100 aa),
32.4%identity in 74 aa overlap"
/codon_start=1
```

Query Match	60 68:	Score 129.6:	DB 1:	Length 19770:
Best Local Similarity	79.08:	Pred. No. 1e-12:		
Matches 169:	Conservative 0:	Mismatches 39:	Indels 6:	Gaps 1:
QY 1	TCAAGGAGAAGCGCTAGCATCTGCGCGCGTGGTGGTACAGGTGAACAAGCTGG 60			
Db 10758	TCAAGGAGAAGCGCTAGCATCTGCGCGCGTGGTGGTACAGGTGAACAAGCTGG 10817			
QY 61	GCTGGGCGGACCAATCCGGCTCAGGTGACACACACCTCAGCGAGGAAGAGCTG 120			
Db 10818	GCTGGGCGGACCAATCCGGCTCAGGTGACACACACCTCAGCGAGGAAGAGCTG 10871			
QY 121	TGCGCACATGAGTACTGTGGTGGCTGACAGAGGGCCAGACAGATGACCGCCCG 180			
Db 10872	TGCGCACATGAGTACTGTGGTGGCTGACAGAGGGCCAGACAGATGACCGCCCG 10931			
QY 181	GGCGGCTGAGAGTCCGCGTGGATGTGAGCAGCAT 214			

Wed Nov 13 13:45:44 2002

us-09-697-123b-21.rge

Page 11

Db 10932 GCGGCTCGAGTGCCTGGGAAACGACGACAT 10965

Search completed: November 13, 2002, 01:27:35
Job time : 679.07 secs

/organism="Mycobacterium smegmatis"
/strain="mc2155"
/db_xref="taxon:1772"
/clone_lib="Mycobacterium anaerobic stationary phase
library"
/lab_host="E. coli XL1-Blue MRF"
/note="Vector: lambda ZAP II; Strata II were disrupted using
a Mini bead beater (Biospec). RNA was isolated using the
RNeasy protocol (Qiagen). Purified RNA was subjected to 2
rounds of digestion with RNase-free DNase I (Promega).
DNase I was heat-inactivated at 75°C for 5 min. and
removed by using RNeasy columns followed by phenol
extraction and ethanol precipitation. The RNA
preparations were confirmed to be free of genomic DNA
contamination by carrying out PCR and RT-PCR using the
Access kit (Promega) and primers specific for the
histone-like protein gene hlp (Lee et al., 1998). cDNA
was synthesized using random hexamer primers (Promega)
and Stratagene's cDNA synthesis kit. cDNA fragments were
ligated into lambda ZAP II vector and packaged in vitro
using Stratagene reagents."

BASE COUNT 230 a 395 c 405 g 252 t
ORIGIN

Query Match 83.8%; Score 174.4; DB 9; Length 1282;
Best Local Similarity 89.9%; Pred. No. 1.1e-28;
Matches 187; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 TCACGAGAGAGCGCTACGACCTGGCCGCGCCGCTACAGGTACACAGAGCTGG 60
DB 236 TCACGAGAGAGCGCTACGACCTGGCCGCGCTTACAGGTACACAGAGCTGG 177
QY 61 GCCTGACGCGCGCCGCGATCAGCTGTCGACTGACCGAGAGAGAGCTGTCGCCA 120
DB 176 GCCTGACGCGCGCGCGATCAGCTGTCGACTGACCGAGAGAGAGCTGTCGCCA 117
QY 121 CCATGAGTACTGCTGGCGCTGCACGAGCGCCAGACAGATGACCGTCCCGCGCGC 180
DB 116 CCATGAGTACTGCTGGCGCTGCACGAGCGGTGACACTTGATGACCTCCCGGTGGCG 57
QY 181 TCGAGTCCCGGTGAGGTGACGACAT 208
DB 56 TCGAGTCCCGGTGAGGTGACGACAT 29

RESULT 2
BI995346
LOCUS 568 bp mRNA linear EST 25-OCT-2001
DEFINITION 1031026E07.Y2 C. reinhardtii CC-1690, Stress II (normalised),
Lambda zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION BI995346
VERSION BI995346.1 GI:16430136
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
REFERENCE 1 (bases 1 to 568)
Grossman, A., Chan, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre
P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished (2001)
JOURNAL Contact: Charles Hauser
COMMENT DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu
FEATURES Location/Qualifiers
source 1..568

/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalised
)", Lambda ZAP II"
/note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - N03- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Wells et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
ZAP II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda ZAP clones by superinfection with Exsist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."

BASE COUNT 80 a 206 c 185 g 97 t
ORIGIN

Query Match 24.1%; Score 50.2; DB 13; Length 568;
Best Local Similarity 52.7%; Pred. No. 0.16;
Matches 109; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 2 CAAGAGAGAGCGCTACGACCTGGCCGCGCGCCGCTACAGGTACACAGAGCTGG 61
DB 251 CAAGAGAGAGCGCTTCCAGAGTGTGACGCGCTTGAACAGGTGATGAGAGTGGCG 310
QY 62 CCTGACGCGCGCCGCGATCAGCTGTCGACTGACCGAGAGAGAGCTGTCGCCA 121
DB 311 CGTGGCTGACGAGGAGCAGCGCGCGCGCTCTCCGCGCGCTCTACACATGGCGCC 370
QY 122 CATGAGTACTGCTGGCGCTGCACGAGCGCCAGACAGATGACCGTCCCGCGCGCT 181
DB 371 CGCGCGCGAGCCCGCGCTGTTCCTCCCGCGCGCGAGGTGTTCGCTCCACACTGGCG 430
QY 182 CGAGTCCCGGTGAGGTGACGACAT 208
DB 431 CGACAGCGCGGTGGCGCTGTTCGACAT 457

RESULT 3
AO687771/c
LOCUS 877 bp DNA linear GSS 01-JUL-1999
DEFINITION nbxb0075L01f CUGI Rice BAC Library Oryza sativa genomic clone
nbxb0075L01f, DNA sequence.
ACCESSION AO687771
VERSION AO687771.1 GI:5328939
KEYWORDS GSS.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eurharididae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 877)
Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
JOURNAL Contact: Ming RA
COMMENT Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCACTATAGG
Class: BAC ends
High quality sequence stop: 85.
FEATURES Location/Qualifiers


```

/tissue_type="Roots"
/dev_stage="Full tillering"
/lab_host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid
bluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Hydropenic
plants grown to full tillering stage were treated with 150
mM NaCl for either 12 hours or 7 days. Root tissues of the
plants subjected to both types of treatment were collected
separately at University of California, Davis (E. Akhnov
and K. Deal in J. Dvorak's lab). Total RNA was prepared
separately from the two samples (12h and 7day treatments),
and equal amount of RNA was then pooled. PolyA RNA was
purified from the pooled RNA, a cDNA library was made, and
the cDNA clones were in vivo excised to give Bluescript
SK(-) phagemids in J. Dvorak's lab (E. Akhnov, J. Dvorak)
at the University of California, Davis. Colony plating,
plasmid DNA preparations and DNA sequencing were performed
in the OD Anderson lab (all other authors)."
BASE COUNT      92 a      198 c      130 g      142 t
ORIGIN
Query Match      21.3%; Score 44.2; DB 14; Length 562;
Best Local Similarity 54.7%; Pred. No. 3.3;
Matches 88; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
OY 25 CCGCGGTGGCGCCGTACAGAGTCACAGAGTCGGCTGAACGCCGCCACCGCATCA 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 160 CCGTCATGCTCTCCGCCCGCCGTCACATCTTCAAGAGCCCTCAACATCTCAAGCCGCCCA 219
OY 85 CGTCGCTACCTCTGACCGAGAGACGTCGCGCCACCATCATGAGTCTGTCGCGCTGC 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 220 CCGGCTGGAAGACGACGACGTCATCTGCGCAGCCCTCGCGCGCTGCTCTCC 279
OY 145 ACGAGGCCACACACGATGACCGTCCCGCGCGCTGAG 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 280 TCGAGGCCATCATCGTGGCGCTGCTGCGCGCCGCCAAG 320

RESULT 6
LOCUS      BE227740      370 bp      mRNA      linear      EST 06-JUL-2000
DEFINITION      894034G12.y1 C. reinhardtii CC-1690, normalized, Lambda zap II
ACCESSION      BE227740
VERSION        BE227740.1 GI:8932979
KEYWORDS
SOURCE
ORGANISM      Chlamydomonas reinhardtii.
                Chlamydomonas reinhardtii.
                Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
                Chlamydomonadaceae; Chlamydomonas.
REFERENCE
AUTHORS      Grossman, A., Davies, J., Federspiel, N., Harris, E., Lefebvre, P.,
                McDermost, J.P., Silflow, C., Stern, D., and Surzycki, R.
TITLE        Analysis of the Chlamydomonas reinhardtii Genome: A Model,
                Unicellular System for Analyzing Gene Function and Regulation in
                Vascular Plants; project phase 2
JOURNAL
COMMENT      Unpublished (2000)
                Contact: Elizabeth H. Harris
                DCMB Box 91000
                Duke University
                Durham, NC 27708-1000, USA
                Tel: 919 613 8164
                Fax: 919 613 8177
                Email: chlamy@duke.edu.
FEATURES
SOURCE
    1..370
    /organism="Chlamydomonas reinhardtii"
    /strain="CC-1690 wild type mt+ 21gr"
    /db_xref="taxon:3055"
    /clone_lib="C. reinhardtii CC-1690, normalized, Lambda zap
    II"
    /note="Vector: Bluescript II SK(-); Site_1: EcoRI; Site_2:
    XhoI; This library, constructed by John Davies and Jeffrey

```

```

McDermost, combines cDNAs from CC-1690 cells grown to
mid-log phase in YAP (acetate-containing) medium in the
light, YAP medium in the dark, HS (minimal) medium in
ambient levels of CO2 and HS medium bubbled with 5% CO2.
PolyA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites.
bluescript II SK- plasmids were excised from the lambda
zap clones by superinfection with Exassist (Stratagene)
phase. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."
BASE COUNT      68 a      123 c      119 g      60 t
ORIGIN
Query Match      20.9%; Score 43.4; DB 10; Length 370;
Best Local Similarity 50.7%; Pred. No. 4.8;
Matches 104; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
OY 2 CAAAGAGAGCGCTACGACCTGGCCGCCGTGAGCCCTACAGGTCACAGAGCTGGG 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 113 CAGCCACACACCCCGCGCGCCGAGAACGACTTCGCGCATCAAGTTCACATCAACAGCGG 172
OY 62 CTTGAAGCGCGCGCCAGCCGATCAGTCGTCGACTCTGACCGAGAGAGAGCTGTGCCAC 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 173 CGAGCCCGCGCGCGAGCGCATCCGCAAGATCTTTGGCGAGACCTCCAGAGTGGCCAC 232
OY 122 CATCGAGTACTGTCGCTGCGCTGCAGGAGCCAGACAGATGACCTCCCGCGCGCT 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 233 GCTCAACATGCGCGACATCCCGCAGCTGCGTGAAGAGTGGCGGTACCAAGTTCCG 292
OY 182 CGAGGTCCCGGTGAGGTGAGCAGC 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 293 CGAGTTGAGGTGAGGTGAGTGGAC 317

RESULT 7
LOCUS      B0767689      391 bp      mRNA      linear      EST 26-JUL-2002
DEFINITION      EBr008_SQ009_H13.R root, 3 week, drought-stressed, cv Optic, EBr008
ACCESSION      B0767689
VERSION        B0767689.1 GI:21976163
KEYWORDS
SOURCE
ORGANISM      Hordeum vulgare.
                Hordeum vulgare.
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
                ; Triticeae; Hordeum.
REFERENCE
AUTHORS      Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
                Ramsay, L., Machray, G., Marshall, D.F.M., and Maugh, R.
TITLE        Development of Barley Transcriptome Resources
JOURNAL
COMMENT      Unpublished (2001)
                Contact: Maugh R. Marshall DF
                Genome Dynamics/Computational Biology
                Scottish Crop Research Institute
                Invergowrie, Dundee, DD2 5DA, Scotland, UK
                Tel: 00 44 1382 562731
                Fax: 00 44 1382 562426
                Email: estescri.sari.ac.uk.
FEATURES
SOURCE
    1..391
    /organism="Hordeum vulgare"
    /cultivar="Optic"
    /db_xref="taxon:4513"
    /clone_lib="EBr008_SQ009_H13"
    /clone_lib="root, 3 week, drought-stressed, cv Optic,
    EBr008"
    /tissue_type="root"
    /dev_stage="3 week"
    /lab_host="DH10B"
    /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
    Non-normalised library, directionally cloned into pSPORT1.

```


Query Match	20.5%	Score 42.6	DB 17	Length 925
Best Local Similarity	14.3%	Pred. No. 7.7		
Matches 29	Conservative 101	Mismatches 73	Indels 0	Gaps
2	CAAGGACAAACGCTACGACCTGGCCGCCGTGGCCGCTTCAAGGTCAACAGAAAGCTGG	61		
774	SASSSRSSGGGAGGGAGGASRRSSSSSSASAGSVSSASSSSSSCSSSVSSASMS	715		
62	CTGGAAGCGCCGCCACCCGATCACTGTCGACTTGACCGAGGAAGCGTGTGCCAC	121		
714	CSBSSSSASASSSSSSSSASCSGCCCTTSSCSSTASMSAARSSSSSSSSCSSSM	65		
122	CATCGAGTACCTGTGGCGCTCCACGAGGGCCAGAACCAAGTGAACGTTCCCGCGGCGT	181		

18	GACCTGGCCCCGGGGGGCCGCTACAAAGGTCACAAAGAACTGGGCTCGAAACGCCGCCAG	77
19		78
92	GCCATGCGCCGACAGCGCCCTCCGCCCGGGGCCACAAGAGCTCTCCCCACAGGTCCTCCCTCCG	15
93		16
94		17
95		18
96		19
97		20
98	CCGATTCAGCTGCTGACTCTGCAGCGAGAAAGAGCGTCGTCGCGCACCATCCGAGTACCTGGTG	13
99		14
152	CTCCCTCTCTCTATCCAGCGACGCGCGCGGCTCTCTCTCCGCGCACCGCGCCCTCGCG	21
153		22
154		23
155		24
156		25
157		26
158	CGCCTGCAACGAGGGCCAGACACGATGACCGTCCCGGCGCGTCGAGGTCGGCGGACAG	197
159		198
212	CTCCGCGCCCTCTCGTGGCTCTCCAGCGCGCTCTCAGGCGGGCGCTTCGAGTGGGCCCAAGAG	277
213		278
214		279
215		280
216		281
217		282
218		283
219		284
220		285
221		286
222		287
223		288
224		289
225		290
226		291
227		292
228		293
229		294
230		295
231		296
232		297
233		298
234		299
235		300
236		301
237		302
238		303
239		304
240		305
241		306
242		307
243		308
244		309
245		310
246		311
247		312
248		313
249		314
250		315
251		316
252		317
253		318
254		319
255		320
256		321
257		322
258		323
259		324
260		325
261		326
262		327
263		328
264		329
265		330
266		331
267		332
268		333
269		334
270		335
271		336
272	GAGGAGA	279
273		280
274		281
275		282
276		283
277		284
278		285
279		286
280		287
281		288
282		289
283		290
284		291
285		292
286		293
287		294
288		295
289		296
290		297
291		298
292		299
293		300
294		301
295		302
296		303
297		304
298		305
299		306
300		307
301		308
302		309
303		310
304		311
305		312
306		313
307		314
308		315
309		316
310		317
311		318
312		319
31		

RESULT 12

QY	1	TCAGGGAAGAGCGCTACGACCTGCGCCCGCTGGCGCGCTTACAGTCAACAGAACGCTTG	60
Db	104	TCACGCGGACATACGACCTCTGCTCTCCGCGCACCGCGGCTCCACATATCTTCTGGCGAGGTGCAC	163
QY	61	GCCTGAACGCGCGCCAGACCGCATCAGCGCTGCGACTGTGACCTGACCCGAGGAAGACGTCTTGCCA	120
Db	164	TCCCGCGGACACATCTGCTCTCACCTCCCGGCGGTGGCGAAGAACATCTACCCCGG	223
QY	121	CCATCGAGTACTGTGTGCGCGCTGACAGGCGCCAGACACGATGACCTGCGCGCGCGG	180
Db	224	TCATCGACCCACCCGCGAGGCGCTGGCGACGTGGAGGACACATCACACCCGATGGGCTCTA	283
QY	181	TCGAGGTCGCGGTCGAGGTGAGACGACA	207
Db	284	ACGTCTGCGCGCCGAGCGCTTTCGAGACGA	310
RESULT 13			
LOCUS	AY108449	1072 bp	mRNA linear HTC 25-MAY-2002
DEFINITION	Zea mays PCO131734 mRNA sequence.		
ACCESSION	AY108449		
VERSION	AY108449.1	GI:21211527	
KEYWORDS	HTC.		
SOURCE	Zea mays.		
ORGANISM	Zea mays.		
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 1072) Halhey,C.F., Dolan,M., Miso,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.		
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes		
JOURNAL	Unpublished (2002)		
REFERENCE	2 (bases 1 to 1072)		
AUTHORS	Coe,E.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
FEATURES	Location/Qualifiers		
source	1. 1072 /organism="Zea mays" /db_xref="MaizeDB:637686" /db_xref="taxon:4577" /clone="PCO131734" /clone_lib="Maize Mapping Project/DuPont Consensus Library" /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"		
BASE COUNT	216 a 348 c 318 g 190 t		
ORIGIN			
Query Match	20.3%;	Score 42.2;	DB 11; Length 1072;
Best Local Similarity	51.9%;	Prod. No. 9.6;	
Matches	95; Conservative	0; Mismatches	88; Indels 0; Gaps 0;
QY	21	CTGGCCCGCGGTGGCGCGCTACAGGTACACAGAACGTGGCGCTGAAAGCGCGGACGCC	80
Db	300	CTGATCAAGCTAGACGCGCGCGCGGCGAGCGCTGAGAGTGGCGCTTTCACAGGCGTACTCG	359
QY	81	ATCAGTCGTGCGACTGTGACCGAGGAAGACGTGCTGCCACCATCGAGTACTGTGTGCGG	140
Db	360	CTGTGGCGAGGCTGTGGCGCTGCGCGCGGAGCGCAAGGTCTATCGCATTCGACGTGAGC	419
QY	141	CTGCAGCAGGCGCCAGACACGATGACGCTCCCGGGCGGCTGAGGTTCCCGGTGAGGTG	200
Db	420	CGGAGTACTACGACATCGGCGCCCGCTTATCGACGCGCGCGGTGCGGCGCAAGGTG	479

OY 201 GAC 203
Db 480 GAC 482

RESULT 14 BG315577

DEFINITION BG315577 453 bp mRNA linear EST 28-NOV-2001
SAB1H09.Y1 Gm-cl032 glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl032-3450 5' similar to TR:09SS31 09SS31 CALMODULIN-LIKE
PROTEIN. ; mRNA sequence.

ACCESSION BG315577
VERSION BG315577.1 GI:13125034
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 453)
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

TITLE
JOURNAL
COMMENT

FEATURES
SOURCE
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccuteresgen.com
High quality sequence stop: 180.

Location/Qualifiers

1..453
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl032-3450"
/clone_1lb="Gm-cl032"
/tissue_type="Cotyledons of 8-day-old 'Williams' seedlings"
/lab_host="DH10B"

/note="Vector: pBluescript II XR; Site.1: EcoRI; Site.2: XhoI; This cDNA library was constructed from mRNA isolated from cotyledons of 8-day-old 'Williams' seedlings which were propagated on paper towels with distilled water for 3 days (etiolated), then greenhouse grown for 5 days in potting soil. The cotyledons were flash-frozen in liquid nitrogen. StrataGene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand ligated cDNA was hememethylated. A modification of StrataGene's first-strand synthesis primer was used. An 'anchor' nucleotide (V-A, C or G) was added to the 3' end of the primer [GAGACAGACAGACAGACAGACTGTCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (400u/ul); all XhoI sites in the cDNA would be protected by their hememethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GIBCOBRL Life Technologies' cDNA Size

Fractionation column. The column eluent was then precipitated, redissolved, and ligated into StrataGene's pBluescript II XR predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated by StrataGene). 100% of the white and 87.5% of the blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=28 and 8 respectively). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

BASE COUNT 97 a 157 c 120 g 78 t 1 others
ORIGIN

Query Match 20.2%; Score 42; DB 12; Length 453;
Best Local Similarity 62.3%; Pred. No 9.9;
Matches 66; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
OY 101 CGAGAACAGCTGCTCGCCACCATGACATCGTGGTCCGCTCAGAGGCCAGACAC 160
DB 282 CGACACGACGCGCTGTCACGCGCCAGACCTCGAGCCCTCTCACGTCCTTGCCGC 341
OY 161 GATGACCTGCGCCGCGCGCGTGCAGTGCCTGTCAGAGTGCAGAC 206
DB 342 GTCCGCGTGGCCCGACGACGCTGCTGCTGCGGAGGTCGAC 387

RESULT 15
LOCUS BM326079 468 bp mRNA linear EST 04-JAN-2002
DEFINITION P1C1_65_E06.b1 A002 Pathogen-Infected compatible 1 (P1C1) Sorghum
bicolor cDNA, mRNA sequence.
ACCESSION BM326079
VERSION BM326079.1 GI:18065216
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor

REFERENCE 1 (bases 1 to 468)
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.

TITLE
JOURNAL
COMMENT
Unpublished (2002)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmp@prattuga.edu

Sequences have been trimmed to exclude polyA, vector, and regions below phred quality 16. The threshold for highest quality sequence is 20. Three-prime sequences, which are obtained with polyTmix or T7 sequencing primer, are presented as the reverse complement.
Seq primer: JEN REV
High quality sequence stop: 416
POLYA=NO.

FEATURES
SOURCE
Location/Qualifiers

1..468
/organism="Sorghum bicolor"
/cultivar="Brx623"
/db_xref="taxon:4558"
/clone_1lb="Pathogen-Infected compatible 1 (P1C1)"
/tissue_type="Leaves"
/dev_stage="4-week-old seedlings infected with Colletotrichum graminicola"

/note="Vector: pBluescript II SK(-) from Lambda Zap II; Site.1: XhoI; Site.2: EcoRI; Four-week-old sorghum seedlings were sprayed with spore suspension prepared from 3-week-old FRM421, a sorghum isolate of the anthracnose pathogen Colletotrichum graminicola. Inoculated plants were kept in a 25 C dark growth chamber with 100% relative

humidity for 24 hr, followed by 12/12 hr of light/dark cycle at 25 C with 90% relative humidity for another 24 hr. All leaves were harvested and quick frozen with liquid nitrogen and stored in a -80 C freezer. The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision. WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen.

BASE COUNT 111 a 150 c 136 g 71 t
ORIGIN

Query Match 20.2%: Score 42; DB 13; Length 468;
Best Local Similarity 51.6%: Pred. No. 9.9;
Matches 96; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

OY 22 TGGCCCGCGTGGGCGCCCTACAAAGTCACAGAGCTGGGCTGAACGCCGCCAGCCGA 81
DB 162 TGGACTACTTCACTACGACTACGAGTTGCGCGAGCCGCCGCGTGAAGCCTGGAGA 221
OY 82 TCACGTCTGACCTGACCGAGAGAGGTGTCGCCACCATCGAGTACTGCTGCGCC 141
DB 222 ACACGCGAGCCGAGCGCCACCAAGCGGACTTCGGGAGAGACACTACTTCGTCGGGACC 281
OY 142 TGCACGAGGGCCAGACACGATGACCGTCCCGCGGCGGTGAGAGTCCCGGTGAGGTGG 201
DB 282 ACGGTGGGTTGAGTCATCATCCACGCCATCGGCGAGCTCTACCTCTCCACGACGCCA 341
OY 202 ACGACA 207
DB 342 ACGGCA 347

Search completed: November 13, 2002, 04:00:27
Job time : 1122.63 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 : Search time 21.1607 Seconds
(without alignments)
3487.380 Million cell updates/sec

Title: US-09-697-123b-19

Sequence: 1 tcaagagagaagcgtacgacac.....ccgctcaggtgacgacacat 208

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88	42.3	5096	10 US-09-984-711-5	Sequence 5, Appl1
2	88	42.3	5099	9 US-10-075-460-5	Sequence 5, Appl1
3	88	42.3	5099	10 US-09-887-052-1	Sequence 1, Appl1
4	88	42.3	5099	10 US-09-887-052-3	Sequence 3, Appl1
5	88	42.3	5099	10 US-09-887-052-5	Sequence 5, Appl1
6	42.2	20.3	1057	10 US-09-452-239-5	Sequence 5, Appl1
7	42	20.2	2693	10 US-09-880-107-3678	Sequence 3678, Ap
8	41.6	20.0	492	10 US-09-815-242-7946	Sequence 7946, Ap
9	40.4	19.4	392	10 US-09-815-242-7946	Sequence 7946, Ap
10	38.6	18.6	1050	12 US-10-060-857-7	Sequence 7, Appl1
11	38	18.3	1131	10 US-09-934-778-1	Sequence 1, Appl1
12	38	18.3	1467	10 US-09-887-576-777	Sequence 777, App
13	37.8	18.2	484	10 US-09-854-731-1	Sequence 1, Appl1
14	37.8	18.2	1164	10 US-09-815-242-7696	Sequence 7696, Ap
15	37.4	18.0	1535	12 US-10-052-586-545	Sequence 545, App
16	37	17.8	569	10 US-09-864-761-15816	Sequence 15816, A
17	37	17.8	1640	10 US-09-925-300-229	Sequence 229, App
18	36.8	17.7	471	10 US-09-815-242-7717	Sequence 7717, Ap
19	36.8	17.7	1266	10 US-09-815-242-7920	Sequence 7920, Ap

20	36.4	17.5	278	10 US-09-923-876-2599	Sequence 2599, Ap
21	36.4	17.5	714	10 US-09-861-289-17	Sequence 17, Appl1
22	36.4	17.5	13613	10 US-09-861-289-3	Sequence 3, Appl1
23	36.2	17.4	1350	10 US-09-815-242-7965	Sequence 7965, Ap
24	36	17.3	4074	10 US-09-815-242-7899	Sequence 7899, Ap
25	35.6	17.1	267	10 US-09-923-876-1232	Sequence 1232, Ap
26	35.6	17.1	1162	12 US-10-027-450-21	Sequence 21, Appl1
27	35.6	17.1	2010	10 US-10-032-717-9	Sequence 9, Appl1
28	35.4	17.0	1029	10 US-09-815-242-8006	Sequence 8006, Ap
29	35.4	17.0	1822	8 US-08-822-186-1	Sequence 1, Appl1
30	35.4	17.0	1822	8 US-08-937-755-1	Sequence 1, Appl1
31	35.4	17.0	1822	9 US-09-982-543A-9	Sequence 9, Appl1
32	35.4	17.0	1822	10 US-09-045-331-1	Sequence 1, Appl1
33	35.4	17.0	1822	10 US-09-828-607-1	Sequence 1824, Ap
34	35.4	17.0	1878	10 US-09-954-456-1824	Sequence 4086, Ap
35	35.2	16.9	1359	10 US-09-815-242-4086	Sequence 5, Appl1
36	35.2	16.9	2825	10 US-09-952-677-5	Sequence 4952, Ap
37	35	16.8	274	10 US-09-923-876-4952	Sequence 7971, Ap
38	35	16.8	828	10 US-09-815-242-7971	Sequence 3856, Ap
39	34.8	16.7	409	10 US-09-864-761-32324	Sequence 7935, Ap
40	34.8	16.7	409	10 US-09-878-574-3856	Sequence 109, App
41	34.8	16.7	1665	10 US-09-815-242-7995	Sequence 111, App
42	34.6	16.6	704	12 US-10-062-254-109	Sequence 7738, Ap
43	34.6	16.6	752	12 US-10-062-254-111	Sequence 2873, Ap
44	34.6	16.6	1437	10 US-09-815-242-7738	
45	34.4	16.5	261	10 US-09-923-876-2873	

ALIGNMENTS

RESULT 1
US-09-984-711-5
Sequence 5, Application US/09984711
Patent No. US20020119549A1

GENERAL INFORMATION:
APPLICANT: MOECKEL, Bettina
APPLICANT: BATHE, Brigitte
APPLICANT: STEPHAN, Hans
APPLICANT: KREUTER, Caroline
APPLICANT: HERMANN, Thomas
APPLICANT: PEPPERLE, Walter
APPLICANT: BINDER, Michael
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE
FILE REFERENCE: 204209USO
CURRENT APPLICATION NUMBER: US/09/984,711
CURRENT FILING DATE: 2001-10-31
PRIOR APPLICATION NUMBER: DE10108230.9
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin version 3.1
SEQ ID NO 5
LENGTH: 5096
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: CDS
LOCATION: (702)..(4196)
OTHER INFORMATION:
US-09-984-711-5

Query Match 42.3%; Score 88; DB 10; Length 5096;
Best local Similarity 68.5%; Pred. No. 8.4e-14;

Matches 137; Conservative 0; Mismatches 60; Indels 3; Gaps 1;

QY 9 AAGCGTAGACCTGGCCGCGCTGACGATCAAGATGAGCTGGCGCTGAAC 68
|||||
DB 1578 AAGCGTAGACCTGGCGCTGCGCTGCTTACAAAGATCAACCCAGATCGGCGCTTGT 1637
|||||
QY 69 GCGCGCAAGCCGATACGTCGACGCTGACGAGGAAGAGCTGCGCCACATCGAG 128
|||||
DB 1638 GCGGACCAAGATGCTTTATGATGACTTCTACTGAAGAGCATCGGACACCATCGAG 1694
|||||

RESULT 6
US-09-452-239-5
Sequence 5, Application US/09/452239
Patent No. US20020081693A1
GENERAL INFORMATION:
APPLICANT: Rafalski, Antoni J.
APPLICANT: Feder, Gary M.
APPLICANT: Cahoon, Rebecca E.
TITLE OF INVENTION: Plant Catefeyl-CoA O-Methyltransferase
FILE REFERENCE: BJ1284 US NA
CURRENT APPLICATION NUMBER: US/09/452,239
CURRENT FILING DATE: 1999-12-01
EARLIER APPLICATION NUMBER: 60/110,594

	Query Match	20.2%	Score 42:	DB 10:	Length 2693:
	Best Local Similarity	53.7%	Pred. No. 0.033:		
	Matches 87:	Conservative	0:	Mismatches 75:	Indels 0:
				Gaps	0:
QY	36	CGCTACAGGTCAACAGAGCTGGGCTCTGACGGCGGCACGGGATTCACGTGTCACCT	95		
Db	276	CTCTACAAAGCTCTACCAACAAAGCTGGGTACGGCTTGTCTTACCTGTAATGGCGGGCTGGCC	217		
QY	96	CTGACCGAGGAACAGCTCGTCGGCACACATCGAGTACCTGTGGCCCTTGACAGAGGGCCAG	155		
Db	216	GTGTGCTGTGTGGCGGTGACATGCGCGTGTGTGATGTCTGACAGCGCTGTGGTCTGTGGCTGTG	157		
QY	156	ACCAAGATGACCTTCCCCGGCGGCGGTGAGAGTCCCGGTCGAG	197		
Db	156	GCCACACAGTTCTGTCCTCCATCAAGGGCAAGGGCGCGGCCAG	115		

```
RESULT 8
US-09-815-242-7946
; Sequence 7946, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIORITY FILING DATE: 2001-03-21
; PRIORITY FILING DATE: 2000-03-21
; PRIORITY FILING DATE: 2000-03-21
; PRIORITY FILING DATE: 2000-05-23
; PRIORITY FILING DATE: 2000-05-23
; PRIORITY FILING DATE: 2000-05-23
; PRIORITY FILING DATE: 2000-05-26
; PRIORITY FILING DATE: 2000-10-23
; PRIORITY FILING DATE: 2000-10-23
; PRIORITY FILING DATE: 2000-11-27
; PRIORITY FILING DATE: 2000-11-27
; PRIORITY FILING DATE: 2000-12-22
; PRIORITY FILING DATE: 2000-12-22
; PRIORITY FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7946
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(492)
US-09-815-242-7946

Query Match          20.0%: Score 41.6: DB 10: Length 492:
Best Local Similarity 52.3%: Pred. No. 0.026:
Matches 92: Conservative 0: Mismatches 84: Indels 0: Gaps 0:

OY 21 CTGGCCCGCGTGGCGGCTGTACAGGTCAACAAGAGCTGGGCTGAAGCGCCGACGCCG 80
    ||| || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 142 CTGACCACCTGTGGGCGACGATGATGATGATGATGATGATGATGATGATGATGATG 201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 81 ATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 202 ATCGAAGTGTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 141 CTGCACGAGGCGCCAGACGATGACCGTCCCGCGGCGCTCGAGGTGCTCCCGGTGCA 196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 CTGCTGAAGTCAAGCGCCAGCGGCGCCAGCGCGCGAGTCAAGGCGCACACCGCA 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
US-09-878-574-3598
; Sequence 3598, Application US/09878574
; Patent No. US2002010548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
```

```
; CURRENT FILING DATE: 2001-12-21
; PRIORITY APPLICATION NUMBER: 09/333,535
; PRIORITY FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 3598
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-008-01-B1-C6
US-09-878-574-3598
```

```
Query Match          19.4%: Score 40.4: DB 10: Length 392:
Best Local Similarity 61.3%: Pred. No. 0.051:
Matches 65: Conservative 0: Mismatches 41: Indels 0: Gaps 0:
```

```
OY 101 CGAGGAAGAGCTGTGCGCCACCATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 33 CGACACGAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 92
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 161 GATGACCTGCTCCCGCGGCGGTGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 93 GTGCGCGTGTCCCGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 10
US-10-060-857-7
; Sequence 7, Application US/10060857
; Patent No. US20020132318A1
; GENERAL INFORMATION:
; APPLICANT: Anticancer, Inc.
; APPLICANT: Zhao, Ming
; APPLICANT: Jlang, Ping
; APPLICANT: Xu, Mingxu
; APPLICANT: Yang, Meng
; TITLE OF INVENTION: IMPROVED FLUORESCENT PROTEINS
; FILE REFERENCE: 31276-20032.00
; CURRENT APPLICATION NUMBER: US/10/060,857
; CURRENT FILING DATE: 2002-01-29
; PRIORITY FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1050
; TYPE: DNA
; ORGANISM: Coral
US-10-060-857-7
```

```
Query Match          18.6%: Score 38.6: DB 12: Length 1050:
Best Local Similarity 49.7%: Pred. No. 0.15:
Matches 98: Conservative 0: Mismatches 99: Indels 0: Gaps 0:
```

```
OY 1 TCAGAGGAAGCGCTTACGACCTGCGCGCGCTGCGCGCTTACAGGTCAACAAGAAAGTGC 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 374 TCGAGGGCGAGGGGAGGGGCGCGCGCTTACGAGGGGCGCAACACCGTGAAGCTGAAGGTGA 433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 61 GCTTGAACCGCGCGCGAGCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 434 CCAAGGGCGCGCGCGCGCGCGCTTTCCTTTCGAGACATCTGTGCCCGCGCTTTCAGTACGCT 493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 121 CCATCGAGTACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 494 CCAAGGTGTACGTGAAGACACCGCGCGGCGGATCCCGGACTACAGAAGAGCTGTCTTCCCG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 181 TCGAGTCCCGGTGAG 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 554 AGGCGTTCAAGTGGAG 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 11
US-09-934-778-1
; Sequence 1, Application US/09934778
; Patent No. US20020106772A1
```

```

US-09-887-576-777
: Sequence 777, Application US/09887576
: Patent No. US20020144047A1
: GENERAL INFORMATION:
: APPLICANT: Budworth, P.
: APPLICANT: Brown, D.
: APPLICANT: Chang, H.
: APPLICANT: Zhu, T.
: APPLICANT: Han, B.
: APPLICANT: Wang, X.
: APPLICANT: Cooper, Bret
: TITLE OF INVENTION: Promoters for regulation of plant expression
: FILE REFERENCE: 160. 001US1
: CURRENT APPLICATION NUMBER: US/09/887,576
: PRIORITY FILING DATE: 2001-06-25
: PRIORITY APPLICATION NUMBER: US 60/213,848
: PRIORITY FILING DATE: 2000-06-23
: PRIORITY APPLICATION NUMBER: US 60/214,087
: PRIORITY FILING DATE: 2000-06-23
: PRIORITY APPLICATION NUMBER: US 60/258,692
: PRIORITY FILING DATE: 2000-12-29
: NUMBER OF SEQ. ID NOS.: 875
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 777
:
: TYPE: DNA
LENGTH: 1467

```

Query Match	18.2%	Score 37.8	DB 10	Length 484
Best Local Similarity	54.9%	Pred. No. 0.23		
Matches	95	Conservative	0	Mismatches 77; Indels 1; Gaps 1
OY	28	GCCTGGCCCTACAGGTCACAAACAACCTGGCTGACGCCGACCGATCAGT	87	
Db	200	GCCTGAGAGCGGCCACGGCAGGATTCCTGTCACATCCATCTGAGGCAACTCGT	259	
OY	88	CGTGACTCTGACCGAGAGAGAGCTGTCGCCACATGAGTACTGCTGGCTGCAGC	147	
Db	260	CACCGACGACGA-CGTGAGAGAGCTCCGCCGGAGATCAATATATACCACTGCGG	318	
OY	148	AGGGCCAGACCAAGATGACCCGCGGGCGGCTGAGAGTCCCGTGCAGGTG	200	

Db 319 GCCACCCCAAGCTGATCTCCATCCGCGCGCTACAGAGAGCGCGCTGCCG 371

RESULT 14

US-09-815-242-7696

Sequence 7696, Application US/09815242
Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Hall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: Prokaryotes
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 7696
LENGTH: 1164
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1164)
US-09-815-242-7696

Query Match 18.28; Score 37.8; DB 10; Length 1164;
Best Local Similarity 50.38; Pred. No. 0.24;
Matches 93; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 23 GGGCCGGGTGGGGCGCTACAGAGTCAACAGAGCTGGGCTGAACGCCGCCGCGAT 82
Db 558 GGGCGATGGTGGCGCGCTCAAGGTTTCCACCAAGCTGAGCTGTAACAGCTGGCCGG 617
QY 83 CACGTCTGCACTGTACCGAGAGAGCGTCTGGCCACCATCGAGTACTGCTGGCGCT 142
Db 618 CATTTGCGACCAACTGATCTCGCGCGGCGCATCGCCAAACCTTCTCGCGCGCGCG 677
QY 143 GCACGAGGCGCAGACGACGATGACCGTCCCGCGCGCGCTGAGTCTCCGCTGAGGTGGA 202
Db 678 ACACAAGCTGGCGAAGTCACTGTACGAGGCGGACCTGTGAGACCGCCAGGCCATCGC 737
QY 203 CGACA 207
Db 738 CGCCA 742

RESULT 15

US-10-052-586-545

Sequence 545, Application US/10052586
Patent No. US20020127584A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: US/10/052,586
CURRENT FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/064466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107

```

1  PRIOR APPLICATION NUMBER: 60/088029
2  PRIOR FILING DATE: 1998-06-04
3  PRIOR APPLICATION NUMBER: 60/088033
4  PRIOR FILING DATE: 1998-06-04
5  PRIOR APPLICATION NUMBER: 60/088167
6  PRIOR FILING DATE: 1998-06-05
7  PRIOR APPLICATION NUMBER: 60/088202
8  PRIOR FILING DATE: 1998-06-05
9  PRIOR APPLICATION NUMBER: 60/088212
10 PRIOR FILING DATE: 1998-06-05
11 PRIOR APPLICATION NUMBER: 60/088217
12 PRIOR FILING DATE: 1998-06-05
13 PRIOR APPLICATION NUMBER: 60/088326
14 PRIOR FILING DATE: 1998-06-04
15 PRIOR APPLICATION NUMBER: 60/088655
16 PRIOR FILING DATE: 1998-06-09
17 PRIOR APPLICATION NUMBER: 60/088722
18 PRIOR FILING DATE: 1998-06-10
19 PRIOR APPLICATION NUMBER: 60/088738
20 PRIOR FILING DATE: 1998-06-10
21 PRIOR APPLICATION NUMBER: 60/088740
22 PRIOR FILING DATE: 1998-06-10
23 PRIOR APPLICATION NUMBER: 60/088811
24 PRIOR FILING DATE: 1998-06-10
25 PRIOR APPLICATION NUMBER: 60/088824
26 PRIOR FILING DATE: 1998-06-10
27 PRIOR APPLICATION NUMBER: 60/088825
28 PRIOR FILING DATE: 1998-06-10
29 PRIOR APPLICATION NUMBER: 60/088826
30 PRIOR FILING DATE: 1998-06-10
31 PRIOR APPLICATION NUMBER: 60/088861
32 PRIOR FILING DATE: 1998-06-11
33 PRIOR APPLICATION NUMBER: 60/088863
34 PRIOR FILING DATE: 1998-06-11
35 PRIOR APPLICATION NUMBER: 60/088876
36 PRIOR FILING DATE: 1998-06-11
37 PRIOR APPLICATION NUMBER: 60/089090
38 PRIOR FILING DATE: 1998-06-12
39 PRIOR APPLICATION NUMBER: 60/089105
40 PRIOR FILING DATE: 1998-06-12
41 PRIOR APPLICATION NUMBER: 60/089512
42 PRIOR FILING DATE: 1998-06-16
43 PRIOR APPLICATION NUMBER: 60/089514
44 PRIOR FILING DATE: 1998-06-16
45 PRIOR APPLICATION NUMBER: 60/089538
46 PRIOR FILING DATE: 1998-06-17
47 PRIOR APPLICATION NUMBER: 60/089598
48 PRIOR FILING DATE: 1998-06-17
49 PRIOR APPLICATION NUMBER: 60/089653
50 PRIOR FILING DATE: 1998-06-17
51 PRIOR APPLICATION NUMBER: 60/089908

```

```

Query Match      18.0% ; Score 37.4; DB 12; Length 1535;
Best Local Similarity 53.0%; Pred.No. 0.31;
Matches 80; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY      1 TCAGAGGAAGCGGTACGACTCTGGCCCCGTCGGCCGTACAGAATCAACAAGATCG 60
          ||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
DB      852 TCTGTGGAGACCCCGTATGCCGTCGCCGGGAAAGAGCCTTGAGAGACCAAGAGATCATAG 911
          ||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

QY      61 GCCGTGAACGCCGGCACCGCATACGCTCTCTCACTCTACGCAGAGAGAACGTCGTGCCA 120
          ||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
DB      912 GCATGCTATTGAGGCACATGAAGTATCATACACAGCATCTGTCTTCCGGCATCGGCTTCA 971
          ||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

QY      121 CCATCGAGTACTCTGTGGCCCTGCACAGAGGG 151
          ||| | |||| | |||| | |||| | |||| | |||| | |||| |
DB      972 CCATCAGCGACGTCTGTGAGATCCACAGGCG 1002
          ||| | |||| | |||| | |||| | |||| | |||| | |||| |

Search completed: November 12, 2002, 16:58:42
Job time : 33.1607 secs
```


	Query Match	80.0%	Score 166.4	DB 5	Length 970
	Best Local Similarity	87.5%	Pred. No. 5.3e-31		
	Matches 182	Conservative 0	Mismatches 26	Indels 0	Gaps 0
QY	1 TCAAGAGAGCGCTACGACCTGCCCCGCGTGGCGCGCTACAGGTACACAAAGACCTGG				60
Db	26 TCAAGGAGAGAGCGCTACGACCTGCCCCGCGTGGCGCTATTAAGGTACACAAAGACCTGG				85
QY	61 GCGTGAACGGCGGCGCCAGCCGATCAGCTGCTGCATCTTGACCGAGAGAGACGTGTCGCA				120
Db	86 GCGTGCATGTGCGGCGAGCCCATACGCTGCTGCAGCGCTGACCGAAGAGAGACGTGTCGCA				145
QY	121 CCATTCGAGTACCTGCTGCCCCCTGCACGAGGCGCCAGACACGATGAGACCTGCCCGGGCGGG				180
Db	146 CCATTCGAAATATGTGTCGCGCTTGCAACGAGGCTCAGACCAAGATGACCGTTCGGGGCGCG				205
QY	181 TCGAGGTCCCGGCTCGAGGTGACGACAT 208				
Db	206 TCGAGGTGCGCGGTGAAACCGACGACAT 233				

Query Match	Best Local Similarity	Score	DB 2	Length
Matches 167	Conservative 0	Mismatches 41	Indels 0	Gaps 0
68.5%	80.3%	142.4		3447
Pred. No. 2.5e-25				
QY	1	TCGAAGGAAGCCGCTACGACCTGGCCCGCGTGGCGGTACAAAGTCAACAAGAGCTGG	60	
Db	809	TCGAAGGAAACCTACGACCTGGCCCGCGTGGCGGTACAAAGTCAACAAGAGCTGG	868	
QY	61	GCCGAAGCGCGGCCACCCGATCAGCTGCTGCATCTTGACCCGAGGAAGAGCTGTCGCCA	120	
Db	869	GCTTGCACGCGGGTGATTGATCAGCTGCTCCACCTGCACCGAAGAGATGCTGCGCCA	928	
QY	121	CCATCAGTACCTGTCGTCCTCTGCACGAGGCGCAGACCATGATACCGTCCCGCGCGGG	180	
Db	929	CCATAGAGTACCTGCTGCTTCGTGCTGATGAGGGTCAGTCGACCAATGATCTCCACGCTGGGG	988	
QY	181	TCGAGGATCCCGGTCGACAGTGCAGCAT	208	
Db	989	TAGAAGTCCAGTGGAAACTGACGCTAT	1016	

RESULT 4
US-09-082-614A-57

Db 472 GAC 474

RESULT 7
US-09-103-840A-1/C
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 20.3%; Score 42.2; DB 4; Length 4411529;
Best Local Similarity 52.6%; Pred. No. 0.13; Mismatches 83; Indels 0; Gaps 0;
Matches 92; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 32 GGGCGCTACAGGTCAACAGAGCTGGGCTGAACGCCGCCGACCGATCAGTCGTC 91
Db 595945 GGAGCGGACCGCGGACACTACCCAGATGCACCTGGAGCGCGGCGACAGATGCTGAT 595886
QY 92 GACTCTGACGAGGAGAGCTGCGCGCCACCATCGAGTACTGTCGCCCTGCAGAGGG 151
Db 595885 CACCGCTACCCATACCACTCGGGCCACCATCGCGCGGCTGCTAACCGCGCGC 595826
QY 152 CCAGACCAAGATGACGCTCCCGCGCGCTGAGGTCGCCGTCGAGTGAGAGCAGC 206
Db 595825 CCTGGGACCGCTCGCGGAGTGCAGCGGAGATATTACCGGAGACTGCTGGC 595771

RESULT 8
US-07-923-692C-5
; Sequence 5, Application US/07923692C
; Patent No. 5316931
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon
; APPLICANT: Dawson, William O.
; APPLICANT: Grantlam, George L.
; APPLICANT: Turpen, Thomas H.
; APPLICANT: Turpen, Ann Myers
; APPLICANT: Garger, Stephen J.
; APPLICANT: Grill, Laurence K.
; TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Limbach & Limbach
; STREET: 2001 Ferry Building
; CITY: San Francisco
; STATE: CAL
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/923,692C
; FILING DATE: 31-JUL-1992
; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 600,244
;; FILING DATE: 22-OCT-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 641,617
;; FILING DATE: 16-JAN-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 310,881
;; FILING DATE: 17-FEB-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 160,766
;; FILING DATE: 26-FEB-1988
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 160,771
;; FILING DATE: 26-FEB-1988
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 347,637
;; FILING DATE: 05-MAY-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 363,138
;; FILING DATE: 08-JUN-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 219,279
;; FILING DATE: 15-JUL-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Halluin, Albert P.
;; REGISTRATION NUMBER: 28,957
;; REFERENCE/DOCKET NUMBER: BIOG-20121
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-433-4150
;; TELEFAX: 415-433-8716
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1450 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHEetical: NO
;; ANTI-SENSE: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Oryza sativa
;; IMMEDIATE SOURCE:
;; CLONE: alpha-amy1ase
;; FEATURE:
;; NAME/KEY: CDS (B) LOCATION: 12..1316
;; LOCATION: 12..1316
US-07-923-692C-5

Query Match 18.7%; Score 38.8; DB 1; Length 1450;
Best Local Similarity 52.5%; Pred. No. 0.45;
Matches 85; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
QY 35 CCGCTACAGGTCAACAGAGCTGGGCTGAACGCCGCCGACCGATCAGTCGTCGAC 94
Db 891 CCGCGGCGAGGAGGCAAGCGCGCGGCTGATCGGGTCTGGCGGCAAGGCGAGCAGC 950
QY 95 TCTGACCGAGGAAGACGTCGTCGCCACCATCGAGTACTGTCGCCCTGCAGAGGGCCA 154
Db 951 CTTCGTGACCAACACAGACCGCGCTCGACGACGACTGTGGCGGTTCCCTCCGCACAA 1010
QY 155 GACCAAGATGACGCTCCCGCGGCTGAGAGTCCCGGTCCA 196
Db 1011 GGTCAATGCAAGGCTACGCTACATCTCAACCCCGGCA 1052

RESULT 9
US-08-184-237-5
; Sequence 5, Application US/08184237
; Patent No. 5589367
; GENERAL INFORMATION:
; APPLICANT: Donson, Jon

```

1  APPLICANT: Dawson, William O.
2  APPLICANT: Grantham, George L.
3  APPLICANT: Turpen, Thomas H.
4  APPLICANT: Turpen, Ann Myers
5  APPLICANT: Garger, Stephen J.
6  APPLICANT: Grill, Laurence K.
7  TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
8  NUMBER OF SEQUENCES: 11
9  CORRESPONDENCE ADDRESS:
10 ADDRESS: Limbach & Limbach
11 STREET: 2001 Ferry Building
12 CITY: San Francisco
13 STATE: CAL
14 ZIP: 94111
15
16 COMPUTER READABLE FORM:
17 MEDIUM TYPE: Floppy disk
18 COMPUTER: IBM PC compatible
19 OPERATING SYSTEM: PC-DOS/MS-DOS
20 SOFTWARE: Patent in Release #1.0, Version #1.25
21 CURRENT APPLICATION DATA:
22 APPLICATION NUMBER: US/08/184,237
23 FILING DATE:
24 CLASSIFICATION: 435
25 PRIOR APPLICATION DATA:
26 APPLICATION NUMBER: US 923,692
27 FILING DATE: 31-JUL-1992
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER: US 600,244
30 FILING DATE: 22-OCT-1990
31 PRIOR APPLICATION DATA:
32 APPLICATION NUMBER: US 641,617
33 FILING DATE: 16-JAN-1991
34 PRIOR APPLICATION DATA:
35 APPLICATION NUMBER: US 310,881
36 FILING DATE: 17-FEB-1989
37 PRIOR APPLICATION DATA:
38 APPLICATION NUMBER: US 160,766
39 FILING DATE: 26-FEB-1988
40 PRIOR APPLICATION DATA:
41 APPLICATION NUMBER: US 160,771
42 FILING DATE: 26-FEB-1988
43 PRIOR APPLICATION DATA:
44 APPLICATION NUMBER: US 347,637
45 FILING DATE: 05-MAY-1989
46 PRIOR APPLICATION DATA:
47 APPLICATION NUMBER: US 363,138
48 FILING DATE: 08-JUN-1989
49 PRIOR APPLICATION DATA:
50 APPLICATION NUMBER: US 219,279
51 FILING DATE: 15-JUL-1988
52 ATTORNEY/AGENT INFORMATION:
53 NAME: Halluin, Albert P.
54 REGISTRATION NUMBER: 28,957
55 REFERENCE/DOCKET NUMBER: BIOG-20121 USA
56 TELECOMMUNICATION INFORMATION:
57 TELEPHONE: 415-433-4150
58 TELEFAX: 415-433-8716
59 INFORMATION FOR SEQ ID NO: 5:
60 SEQUENCE CHARACTERISTICS:
61 LENGTH: 1450 base pairs
62 TYPE: nucleic acid
63 STRANDEDNESS: single
64 TOPOLOGY: linear
65 MOLECULE TYPE: DNA (genomic)
66 HYPOHETICAL: NO
67 ANTI-SENSE: NO
68 ORIGINAL SOURCE:
69 ORGANISM: Oryza sativa
70 IMMEDIATE SOURCE:
71 CLONE: alpha-amylase
72 FEATURE:
73 NAME/KEY: CDS (B) LOCATION: 12..1316
74 LOCATION: 12..1316

```

Query Match	18.7%	Score 38.8	DB 1	Length 1450
Best Local Similarity	52.5%	Pred. No. 0.45		
Matches 85	Conservative 0	Mismatches 77	Indels 0	Gaps 0
QY	35	CCGCTACAAAGTCAACAAAGAGCTGGGCGCTAACGCCGGCCAGCCCATTCACGTCGTGCAC	94	
Db	891	CCGGGGGAGAGACGGCAAGAGCGCCCGGGGATATGTCGGGTGGCGCGGCCCAAGCGGACGAC	950	
QY	95	TCGTACCGAAGAAAGCGTGTGCGCCACCATTGAGTACTGTGTGGCGCTGTGCAGAGGCCA	154	
Db	951	CTTGTGTGACAAACAAAGACGACGCGTCTGACGACGAGACTGTGGGCCGTTTCCCTTCGACAA	1010	
QY	155	GACCAGATGAGCCGTCGCCGGGGGGCGTGTGAGAGTCCCGGTGCA	196	
Db	1011	GGTCATGTCAAGGGCTTAAGCATACATCTTACCCACCCGGGCA	1052	

RESULT 10.
 US-08-482-920-5
 Sequence 5, Application US/08482920
 Patent No. 5866785
 GENERAL INFORMATION:
 APPLICANT: Donson, Jon
 APPLICANT: Dawson, William O.
 APPLICANT: Grantham, George L.
 APPLICANT: Turpen, Thomas H.
 APPLICANT: Turpen, Ann Myers
 APPLICANT: Garger, Stephen J.
 APPLICANT: Grill, Laurence K.
 TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/482,920
 FILING DATE: 07-JUNE-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 184,237
 FILING DATE: 19-JAN-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 600,244
 FILING DATE: 22-OCT-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 641,617
 FILING DATE: 16-JAN-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 310,881
 FILING DATE: 17-FEB-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 160,766
 FILING DATE: 26-FEB-1988
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 160,771
 FILING DATE: 26-FEB-1988
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 347,637
 FILING DATE: 05-MAY-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 363,138
 FILING DATE: 08-JUN-1989
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 219,279


```
; SEQ ID NO 1
; LENGTH: 2742
; TYPE: DNA
; ORGANISM: Pseudorabies virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2742)
US-09-232-468A-1
```

```
Query Match
Best Local Similarity 54.1%; Score 38.8; DB 4; Length 2742;
Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
```

```
OY 62 CCTGAACGCCGGCCAGCATCCTGCTGACCTGACCCAGAGAGAGCTGCGCCAC 121
| 1111 111 11 111111 111 11 1111 11111 11111 11
DB 2484 CATGAAGGCCCTGTACCCCGTCACAGAGAGAGCTCAAGAGAGAGCGCTCGACGAAG 2543
OY 122 CATCGAGTACTGTGTGCGCCCTGCACAGAGGCGCAGACCATGACCTGCCCGGCGCGT 181
| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 2544 CGACCTGGACGAGCGCCAGAGCTGAGACAGGCGCGGAGCATGATCCGGTACATGTCATCGT 2603
OY 182 CGAGGTCCCGGTGCGAGTGGAGACACA 207
| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 2604 GTGGGCCCTCGACGACGAGGACACA 2629
```

```
RESULT 13
5196516-7
; Patent No. 5196516
; APPLICANT: SCHREURS, CHRISTA S.; METTENLEITER, THOMAS C.
; SIMON, ARTHUR J.; LUKAS, NOEMI; RZHA, HANNS J.
; TITLE OF INVENTION: PSEUDORABIES VIRUS VACCINE
; NUMBER OF SEQUENCES: 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/383,833
; FILING DATE: 21-JUL-1989
; SEQ ID NO: 7
; LENGTH: 4897
5196516-7
```

```
Query Match
Best Local Similarity 54.1%; Score 38.8; DB 6; Length 4897;
Matches 79; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
```

```
OY 62 CCTGAACGCCGGCCAGCATCCTGCTGACCTGACCCAGAGAGAGCTGCGCCAC 121
| 1111 111 11 111111 111 11 1111 11111 11111 11
DB 3355 CATGAAGGCCCTGTACCCCGTCACAGAGAGAGCTCAAGAGAGAGCGCTCGACGAGG 3414
OY 122 CATCGAGTACTGTGTGCGCCCTGCACAGAGGCGCAGACCATGACCTGCCCGGCGCGT 181
| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 3415 CGACGTGGACGAGCGCCAGAGCTGAGACAGGCGCGGAGCATGATCCGGTACATGTCATCGT 3474
OY 182 CGAGGTCCCGGTGCGAGTGGAGACACA 207
| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 3475 GTGGGCCCTCGACGACGAGGACACA 3500
```

```
RESULT 14
US-09-029-603-4
; Sequence 4, Application US/09029603
; Patent No. 6210935
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Engel, Natalie
; APPLICANT: Bietenhader, Jurg
; APPLICANT: Toupet, Christine
; APPLICANT: Pospisich, Andreas
; TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters
; FILE REFERENCE: 4-2035/A/PCT
; CURRENT APPLICATION NUMBER: US/09/029,603
; EARLIER FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: PCT/EP96/03643
; EARLIER FILING DATE: 1996-08-19
```

```
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 6085
; TYPE: DNA
; ORGANISM: Streptomyces longisporoflavus
; FEATURE:
; NAME/KEY: misc_RNA
; LOCATION: (378)..(1665)
; OTHER INFORMATION: ORF
; FEATURE:
; NAME/KEY: misc_RNA
; LOCATION: (1747)..(2553)
; OTHER INFORMATION: ORF
; FEATURE:
; NAME/KEY: misc_RNA
; LOCATION: (2593)..(4011)
; OTHER INFORMATION: ORF
; FEATURE:
; NAME/KEY: misc_RNA
; LOCATION: (4013)..(4999)
; OTHER INFORMATION: ORF
; NAME/KEY: misc_RNA
; LOCATION: (5071)..(6085)
; OTHER INFORMATION: ORF
US-09-029-603-4
```

```
Query Match
Best Local Similarity 51.1%; Score 38.8; DB 4; Length 6085;
Matches 91; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
```

```
OY 18 GACCTGGCCCGCGTGGGCGCCCTACAGAGTCAACAAAGCTGGCCTGAGCCGCCAG 77
| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1017 GAGGTGAGACCGCGCGCGCGCGAGACCGGAGCATGCTACCTCTGTCGCGCGC 1076
OY 78 CCGATCACGCTGCTGACTGACCCGAGAGAGAGTGTGCGCCAGCATGAGTACCTGTG 137
| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1077 CGGACACCGGATCACGCTCAGCGGTGAGCGCATGCTGTGCGACGCGTCCATGCTC 1136
OY 138 CGCTGCAGCAGGAGGCGCAGACGATGACCGCTCCCGGCGCGTGTGAGTCCCGTGC 195
| 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DB 1137 ACCGCGCGCCAGACGACACACCAACTGCTGCGCAGGCGGTCTGACCTGCGCG 1194
```

```
RESULT 15
US-07-640-476-6
; Sequence 6, Application US/07640476
; Patent No. 5376536
; GENERAL INFORMATION:
; APPLICANT: QUAX, WILHELMUS
; APPLICANT: LOTTEN, RUDOLF G.M.
; APPLICANT: SCHUURHUIZEN, PAUL W.
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASE ENZYMES AND
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 345 Middlefield Road, suite 200
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/640,476
; FILING DATE: 19910110
```

```

: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Kate H. Murashige
: REGISTRATION NUMBER: 29,959
: REFERENCE/DOCKET NUMBER: 24615-20009.20
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 327-7250
: TELEFAX: (415) 327-2951
: TELEX: 706141
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1164 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Streptomyces murinus
: STRAIN: DSM 40091
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1164
: IDENTIFICATION METHOD: experimental
: OTHER INFORMATION: /EC_number=5.3.1.5
: OTHER INFORMATION: /product= "xylose isomerase (glucose isomerase)"
: OTHER INFORMATION: /evidence= EXPERIMENTAL
: OTHER INFORMATION: /standard_name= "D-xylose ketol isomerase"
: US-07-640-476-6

```

```

Query Match      18.6%; Score 38.6; DB 1; Length 1164;
Best Local Similarity 49.7%; Pred. No. 0.49;
Matches 98; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY      2 CAAGGAGAGCGCTACGACCTGGCGCGCGCTACAGGTCAACAAGAGCTGGG 61
      ||| | | | | | | | | | | | | | | | | | | | | | | | | |
DB      717 CAAGCTTTCGACATCGACCTCAACGGCCACTCCGGCATCAAGTACGACGAGCTGG 776
QY      62 CCTGAACCGCGCGCAGCGATCAGCTGCTGACTGTGACGAGGAGAGCGTGGCGCAC 121
      | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      777 GTTCGGCGCGCGCGACCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 836
QY      122 CATCGACTACCTGGTGGCGCTGACGAGAGGCGCACACCATGACCGTCCCGGCGCGT 181
      ||| | | | | | | | | | | | | | | | | | | | | | | |
DB      837 TTACGAGGGCGCGCGCGACTTCAAGCCGCGCGGAGCCGAGGACTTCGACGCGCGT 896
QY      182 CGAGTCCCGGTGAGG 198
      || | | | | | | |
DB      897 GTGGGCTGCGCGCGG 913

```

Search completed: November 12, 2002, 21:00:00
 Job time : 2703.89 secs

PT diagnosis and identification of many mycobacterial species by restriction fragment length polymorphism -

XX Claim 1: Page 46; 50pp: English.

PS The present sequence for *Mycobacterium fortuitum* rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various *Mycobacterium* species. These rpoB gene fragments can be used
CC in the diagnosis and identification of *Mycobacterium* species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of *Mycobacterium* species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including *M. tuberculosis*, *M. leprae*
CC and non-tuberculous *Mycobacterium* (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC experiment), and can differentiate between many species in a single
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific *Mycobacterium* species.

XX Sequence 208 BP; 43 A; 69 C; 69 G; 27 T; 0 other;

Query Match Best Local Similarity 100.0%; Score 208; DB 22; Length 208;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGAGAACGGCTACGACCTGGCCCGCGCTGACGAGTCAACAGAGCTGG 60
DB 1 TCAAGAGAACGGCTACGACCTGGCCCGCGCTGACGAGTCAACAGAGCTGG 60
QY 61 GCCTGACGCGCGCCGCGATCACTGCTGACGAGTCAACAGAGCTGGCCCA 120
DB 61 GCCTGACGCGCGCCGCGATCACTGCTGACGAGTCAACAGAGCTGGCCCA 120
QY 121 CCATGAGTACTGCTGCTGCGCTGACGAGGCGCCAGACGATGACCGTCCCGGCGCG 180
DB 121 CCATGAGTACTGCTGCTGCGCTGACGAGGCGCCAGACGATGACCGTCCCGGCGCG 180
QY 181 TCGAGTCCCGGCTGAGTGCAGACAT 208
DB 181 TCGAGTCCCGGCTGAGTGCAGACAT 208

RESULT 2

AAS05215 standard; DNA; 208 BP.

XX AAS05215;

DT 07-SEP-2001 (first entry)

DE *Mycobacterium avium* rpoB gene fragment.

XX Non-tuberculous *Mycobacterium*: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.

OS *Mycobacterium avium*.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUM BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
PI

XX WPI: 2001-300520/31.

DR New DNA fragments from the rpoB gene of *Mycobacterium*, useful for
XX diagnosis and identification of many *Mycobacterium* species by
PT restriction fragment length polymorphism

PS Claim 1: Page 44; 50pp: English.

XX The present sequence for *Mycobacterium avium* rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various *Mycobacterium* species. These rpoB gene fragments can be used
CC in the diagnosis and identification of *Mycobacterium* species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of *Mycobacterium* species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including *M. tuberculosis*, *M. leprae*
CC and non-tuberculous *Mycobacterium* (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC experiment), and can differentiate between many species in a single
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific *Mycobacterium* species.

XX Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;

Query Match Best Local Similarity 86.2%; Score 179.2; DB 22; Length 208;
Matches 190; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 TCAAGAGAACGGCTACGACCTGGCCCGCGCTGACGAGTCAACAGAGCTGG 60
DB 1 TCAAGAGAACGGCTACGACCTGGCCCGCGCTGACGAGTCAACAGAGCTGG 60
QY 61 GCCTGACGCGCGCCGCGATCACTGCTGACGAGTCAACAGAGCTGGCCCA 120
DB 61 GCCTGACGCGCGCGATCACTGCTGACGAGTCAACAGAGCTGGCCCA 120
QY 121 CCATGAGTACTGCTGCTGCGCTGACGAGGCGCCAGACGATGACCGTCCCGGCGCG 180
DB 121 CCATGAGTACTGCTGCTGCGCTGACGAGGCGCCAGACGATGACCGTCCCGGCGCG 180
QY 181 TCGAGTCCCGGCTGAGTGCAGACAT 208
DB 181 TCGAGTCCCGGCTGAGTGCAGACAT 208

RESULT 3

AAS05211 standard; DNA; 208 BP.

XX AAS05211;

DT 07-SEP-2001 (first entry)

DE *Mycobacterium marinum* rpoB gene fragment.

XX Non-tuberculous *Mycobacterium*: rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.

OS *Mycobacterium marinum*.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

DE Mycobacterium gordonae type I rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
OS Mycobacterium gordonae type I.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI: 2001-300520/31.
DR
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism -
XX
PS Claim 1: Page 40; 50pp; English.
XX
XX The present sequence for Mycobacterium gordonae type I rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments: isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 46 A; 71 C; 65 G; 26 T; 0 other;
Query Match 81.5%; Score 169.6; DB 22; Length 208;
Best Local Similarity 88.5%; Pred. No. 3,7e-29;
Matches 184; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 1 TCAAGGAGAACCGCTACGACCTGCGCCGCGCTACAGAGTCAACAGAAAGCTGG 60
DB 1 TCAAGGAGAACCGCTACGACCTGCGCCGCGCTACAGAGTCAACAGAAAGCTGG 60
QY 61 GCCTGAACGCGCGCCAGCGATCGTGCATCTGCAGCCGAGGAAGAGCTGTGCCA 120
DB 61 GCCTGAACGCGCGCGATCGCATCAACGCTCAACGCTGAGCGAGAGAGCTGTGCCA 120
QY 121 CCATCGAGTACTGTGTGCGCTGCACGAGGCGCAGACGATGACCTGCCGCGGCGG 180
DB 121 CCATCGAGTACTGTGTGCGCTGCACGAGGCGCAGCAGATGACCTGCCGCGGCGG 180
QY 181 TCGAGGTCCCGTTCGAGAGTGCAGACAT 208
DB 181 CCGAGGTCCCGTTCGAGAGTGCAGACAT 208
RESULT 6
AAS05202

ID AAS05202 standard; DNA: 208 BP.
XX
AC AAS05202;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium gordonae type II rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
OS Mycobacterium gordonae type II.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI: 2001-300520/31.
DR
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism -
XX
PS Claim 1: Page 40; 50pp; English.
XX
XX The present sequence for Mycobacterium gordonae type II rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments: isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 45 A; 62 C; 73 G; 28 T; 0 other;
Query Match 81.5%; Score 169.6; DB 22; Length 208;
Best Local Similarity 88.5%; Pred. No. 3,7e-29;
Matches 184; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 1 TCAAGGAGAACCGCTACGACCTGCGCCGCGCTACAGAGTCAACAGAAAGCTGG 60
DB 1 TCAAGGAGAACCGCTACGACCTGCGCCGCGCTACAGAGTCAACAGAAAGCTGG 60
QY 61 GCCTGAACGCGCGCCAGCGATCGTGCATCTGCAGCGAGGAGAGCTGTGCCA 120
DB 61 GCCTGAACGCGCGCGATCGCATCAACGCTCAACGCTGAGCGAGAGAGCTGTGCCA 120
QY 121 CCATCGAGTACTGTGTGCGCTGCACGAGGCGCAGACGATGACCTGCCGCGGCGG 180
DB 121 CCATCGAGTACTGTGTGCGCTGCACGAGGCGCAGTGCAGTGCAGTGCAGTGCAG 180
QY 181 TCGAGGTCCCGTTCGAGAGTGCAGACAT 208

Db 181 CCGAGGTGCCGTGACAGCCGACGACAT 208

RESULT 7

AA505208

AA505208 standard; DNA: 208 BP.

07-SEP-2001 (first entry)

Mycobacterium kansasii rpoB gene fragment.

Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP; PCR-restriction fragment length polymorphism analysis; ds.

Mycobacterium kansasii.

WO200131061-A1.

03-MAY-2001.

27-OCT-2000; 2000WO-KR01223.

27-OCT-1999; 99KR-0046795.

(ERUM-) ERUME BIOTECH CO LTD.

Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

WPI: 2001-300520/31.

New DNA fragments from the rpoB gene of mycobacteria, useful for diagnosis and identification of many mycobacterial species by restriction fragment length polymorphism

Claim 1; Page 42; 50pp; English.

The present sequence for Mycobacterium kansasii rpoB gene fragment is 1 of 24 rpoB gene fragments (AA505201-AA505224) from various Mycobacterial species. These rpoB gene fragments can be used in the diagnosis and identification of Mycobacterium species using a novel PCR-restriction fragment length polymorphism analysis (PRA) method. The method comprises obtaining a restriction fragment length polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA fragment from the microorganism to be identified and comparing the RFLP patterns from the known rpoB gene fragments with the unidentified fragment. The rpoB gene fragments are useful to identify a wide range of Mycobacterium species, e.g. for diagnosis or to obtain epidemiological and pathogenesis information for selection of appropriate therapies, including M. tuberculosis, M. leprae and non-tuberculous mycobacteria (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and cost effective (only 1 PCR required), and can differentiate between many species in a single experiment, including those difficult to distinguish by usual biochemical tests. Also described are oligonucleotide probes (AA505227-AA505242) for detecting specific Mycobacterial species.

Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;

Query Match 80.8%; Score 168; DB 22; Length 208; Best Local Similarity 88.0%; Pred. No. 8.3e-29;

Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

1 TCACGAGACGAGGCTACACCTGCGCCGCGCTACACAGTCAACAGAAAGCTGG 60

1 TCACGAGAGAGGCTACACCTGCGCCGCGCTACACAGTCAACAGAAAGCTGG 60

61 GCCTGACGCGGCGCCGATCATGCTGACTGACCGAGAGAGAGCTGTCGCCA 120

61 GCCTGACACCAATCATCGATCAACAGACGCTGACCGAAGAAAGCTGTCGCCA 120

Qy 121 CCATGAGTACTGTCGCGCTGACAGAGGCCAGACCATGATGACCTGCCGCGCGC 180
Db 121 CCATCGAGTATCTGTCGCGCTGACAGAGGCCAGACCATGATGACCTGCCGCGCGC 180
Qy 181 TCAGAGTCCCGCTGACAGTGGACGACAT 208
Db 181 TCAGAGTCCCGCTGACAGGACGACAT 208

RESULT 8

AA505217

AA505217 standard; DNA: 208 BP.

07-SEP-2001 (first entry)

Mycobacterium celatum rpoB gene fragment.

Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP; PCR-restriction fragment length polymorphism analysis; ds.

Mycobacterium celatum.

WO200131061-A1.

03-MAY-2001.

27-OCT-2000; 2000WO-KR01223.

27-OCT-1999; 99KR-0046795.

(ERUM-) ERUME BIOTECH CO LTD.

Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

WPI: 2001-300520/31.

New DNA fragments from the rpoB gene of mycobacteria, useful for diagnosis and identification of many mycobacterial species by restriction fragment length polymorphism

Claim 1; Page 45; 50pp; English.

The present sequence for Mycobacterium celatum rpoB gene fragment is 1 of 24 rpoB gene fragments (AA505201-AA505224) from various Mycobacterial species. These rpoB gene fragments can be used in the diagnosis and identification of Mycobacterium species using a novel PCR-restriction fragment length polymorphism analysis (PRA) method. The method comprises obtaining a restriction fragment length polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating, amplifying and digesting the DNA fragment from the microorganism to be identified and comparing the RFLP patterns from the known rpoB gene fragments with the unidentified fragment. The rpoB gene fragments are useful to identify a wide range of Mycobacterium species, e.g. for diagnosis or to obtain epidemiological and pathogenesis information for selection of appropriate therapies, including M. tuberculosis, M. leprae and non-tuberculous mycobacteria (NTM) encountered in subjects infected with human immunodeficiency virus (HIV). Analysis of the rpoB gene fragments is rapid, precise, simple and cost effective (only 1 PCR required), and can differentiate between many species in a single experiment, including those difficult to distinguish by usual biochemical tests. Also described are oligonucleotide probes (AA505227-AA505242) for detecting specific Mycobacterial species.

Sequence 208 BP; 48 A; 71 C; 64 G; 25 T; 0 other;

Query Match 80.8%; Score 168; DB 22; Length 208; Best Local Similarity 88.0%; Pred. No. 8.3e-29;

Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

1 TCACGAGACGAGGCTACACCTGCGCCGCGCTACACAGTCAACAGAAAGCTGG 60

Query Match	Best Local Similarity	80.0%	Score 166.4	DB 22	Length 208
Matches 182	Conservative	0	Mismatches 26	Indels 0	Gaps 0
0Y	1	TCACGAGAACGCGGTACGACCTGGCCCGGTGGCGCGTCAACAGTCAACAGAAAGCTGG	60		
Db	1	TCACGAGAACGCGGTACGACCTGGCCCGGTGGCGCGTCAACAGTCAACAGAAAGCTGG	60		
0Y	61	GGCTGAACGCGCGCGCACCGCATCACTGTCGACTGACCGAGGACAGACGTCGTGGCA	120		
Db	61	GGCTGAACGCGCGCGCACCGCATCACTGTCGACTGACCGAGGACAGACGTCGTGGCA	120		
0Y	121	CCATTCGATACCTTGTGTCGCTGACGAGGCGCCAGACGATGACCTCCCGCGCGCG	180		
Db	121	CCATTCGATACCTTGTGTCGCTGACGAGGCGCCAGACGATGACCTCCCGCGCGCG	180		
0Y	181	TCGAGGTCCGCGCTGAGAGTGCACGACAT	208		
Db	181	TCGAGGTCCGCGCTGAGAGTGCACGACAT	208		
RESULT 10					
0Y	1	TCACGAGAACGCGGTACGACCTGGCCCGGTGGCGCGTCAACAGTCAACAGAAAGCTGG	60		
Db	1	TCACGAGAACGCGGTACGACCTGGCCCGGTGGCGCGTCAACAGTCAACAGAAAGCTGG	60		
0Y	61	GGCTGAACGCGCGCGCACCGCATCACTGTCGACTGACCGAGGACAGACGTCGTGGCA	120		
Db	61	GGCTGAACGCGCGCGCACCGCATCACTGTCGACTGACCGAGGACAGACGTCGTGGCA	120		
0Y	121	CCATTCGATACCTTGTGTCGCTGACGAGGCGCCAGACGATGACCTCCCGCGCGCG	180		
Db	121	CCATTCGATACCTTGTGTCGCTGACGAGGCGCCAGACGATGACCTCCCGCGCGCG	180		
0Y	181	TCGAGGTCCGCGCTGAGAGTGCACGACAT	208		
Db	181	TCGAGGTCCGCGCTGAGAGTGCACGACAT	208		
RESULT 10					
0Y	1	TCACGAGAACGCGGTACGACCTGGCCCGGTGGCGCGTCAACAGTCAACAGAAAGCTGG	60		
Db	1	TCACGAGAACGCGGTACGACCTGGCCCGGTGGCGCGTCAACAGTCAACAGAAAGCTGG	60		
0Y	61	GGCTGAACGCGCGCGCACCGCATCACTGTCGACTGACCGAGGACAGACGTCGTGGCA	120		
Db	61	GGCTGAACGCGCGCGCACCGCATCACTGTCGACTGACCGAGGACAGACGTCGTGGCA	120		
0Y	121	CCATTCGATACCTTGTGTCGCTGACGAGGCGCCAGACGATGACCTCCCGCGCGCG	180		
Db	121	CCATTCGATACCTTGTGTCGCTGACGAGGCGCCAGACGATGACCTCCCGCGCGCG	180		
0Y	181	TCGAGGTCCGCGCTGAGAGTGCACGACAT	208		
Db	181	TCGAGGTCCGCGCTGAGAGTGCACGACAT	208		
RESULT 10					
0Y	1	TCACGAGAACGCGGTACGACCTGGCCCGGTGGCGCGTCAACAGTCAACAGAAAGCTGG	60		
Db	1	TCACGAGAACGCGGTACGACCTGGCCCGGTGGCGCGTCAACAGTCAACAGAAAGCTGG	60		
0Y	61	GGCTGAACGCGCGCGCACCGCATCACTGTCGACTGACCGAGGACAGACGTCGTGGCA	120		
Db	61	GGCTGAACGCGCGCGCACCGCATCACTGTCGACTGACCGAGGACAGACGTCGTGGCA	120		
0Y	121	CCATTCGATACCTTGTGTCGCTGACGAGGCGCCAGACGATGACCTCCCGCGCGCG	180		
Db	121	CCATTCGATACCTTGTGTCGCTGACGAGGCGCCAGACGATGACCTCCCGCGCGCG	180		
0Y	181	TCGAGGTCCGCGCTGAGAGTGCACGACAT	208		
Db	181	TCGAGGTCCGCGCTGAGAGTGCACGACAT	208		
RESULT 10					
0Y	1	TCACGAGAACGCGGTACGACCTGGCCCGGTGGCGCGTCAACAGTCAACAGAAAGCTGG	60		
Db	1	TCACGAGAACGCGGTACGACCTGGCCCGGTGGCGCGTCAACAGTCAACAGAAAGCTGG	60		
0Y	61	GGCTGAACGCGCGCGCACCGCATCACTGTCGACTGACCGAGGACAGACGTCGTGGCA	120		
Db	61	GGCTGAACGCGCGCGCACCGCATCACTGTCGACTGACCGAGGACAGACGTCGTGGCA	120		
0Y	121	CCATTCGATACCTTGTGTCGCTGACGAGGCGCCAGACGATGACCTCCCGCGCGCG	180		
Db	121	CCATTCGATACCTTGTGTCGCTGACGAGGCGCCAGACGATGACCTCCCGCGCGCG	180		
0Y	181	TCGAGGTCCGCGCTGAGAGTGCACGACAT	208		
Db	181	TCGAGGTCCGCGCTGAGAGTGCACGACAT	208		
RESULT 10					
0Y	1	TCACGAGAACGCGGTACGACCTGGCCCGGTGGCGCGTCAACAGTCAACAGAAAGCTGG	60		
Db	1	TCACGAGAACGCGGTACGACCTGGCCCGGTGGCGCGTCAACAGTCAACAGAAAGCTGG	60		
0Y	61	GGCTGAACGCGCGCGCACCGCATCACTGTCGACTGACCGAGGACAGACGTCGTGGCA	120		
Db	61	GGCTGAACGCGCGCGCACCGCATCACTGTCGACTGACCGAGGACAGACGTCGTGGCA	120		
0Y	121	CCATTCGATACCTTGTGTCGCTGACGAGGCGCCAGACGATGACCTCCCGCGCGCG	180		
Db	121	CCATTCGATACCTTGTGTCGCTGACGAGGCGCCAGACGATGACCTCCCGCGCGCG	180		
0Y	181	TCGAGGTCCGCGCTGAGAGTGCACGACAT	208		
Db	181	TCGAGGTCCGCGCTGAGAGTGCACGACAT	208		
RESULT 10					
0Y	1	TCACGAGAACGCGGTACGACCTGGCCCGGTGGCGCGTCAACAGTCAACAGAAAGCTGG	60		
Db	1				

CC required), and can differentiate between many species in a single
 CC experiment, including those difficult to distinguish by usual biochemical
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
 CC detecting specific Mycobacterial species.

SO Sequence 208 BP: 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 80.0%; Score 166.4; DB 22; Length 208;
 Best Local Similarity 87.5%; Pred. No. 1.9e-28;
 Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 1 TCAGAGGAAGCGCTACGACCTGGCCCGCGGTGGCGCTACAGGTCAACAAGAACTGG 60
 DB 1 TCAGAGGAAGCGCTACGACCTGGCCCGCGGTGGCGCTACAGGTCAACAAGAACTGG 60
 OY 61 GCCTGAACCGCGCCGACGATCAGCTGCTGCTACCTGACGAGAGAGAGCTGCGCA 120
 DB 61 GCCTGATGTGCGCGGACGACATCAGCTGCTGACGCTGACGAGAGAAAGAACTGCGCA 120
 OY 121 CCATCGAGTACCTGTGCTGCGCTGACGAGGCGCAGACCATGATGCGCGCGCGCG 180
 DB 121 CCATCGAATATCTGTGCTGCGCTGACGAGGCGTCAACCATGATGCGCTGCGCGCGCG 180
 OY 181 TCGAGGTCCCGGTGAGGTGAGCAGCAT 208
 DB 181 TCGAGGTCCCGGTGAGAAACGACGACAT 208

RESULT 11
 AAS05216
 ID AAS05216 standard; DNA: 208 BP.

AC AAS05216;

DT 07-SEP-2001 (first entry)

DE Mycobacterium bovis rpoB gene fragment.

KM Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
 KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium bovis.

PN M0200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000MO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

PT New DNA fragments from the rpoB gene of mycobacteria, useful for
 PT diagnosis and identification of many mycobacterial species by
 PT restriction fragment length polymorphism

PS Claim 1: Page 45; 50pp; English.

CC The present sequence for Mycobacterium bovis rpoB gene
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05242) from
 CC various Mycobacterial species. These rpoB gene fragments can be used
 CC in the diagnosis and identification of Mycobacterium species using a
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)
 CC method. The method comprises obtaining a restriction fragment length
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
 CC amplifying and digesting the DNA fragment from the microorganism to
 CC be identified and comparing the RFLP patterns from the known rpoB gene
 CC fragments with the unidentified fragment. The rpoB gene fragments

CC are useful to identify a wide range of Mycobacterium species, e.g. for
 CC diagnosis or to obtain epidemiological and pathogenesis information for
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR
 CC required), and can differentiate between many species in a single
 CC experiment, including those difficult to distinguish by usual biochemical
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
 CC detecting specific Mycobacterial species.

SO Sequence 208 BP: 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 80.0%; Score 166.4; DB 22; Length 208;
 Best Local Similarity 87.5%; Pred. No. 1.9e-28;
 Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 1 TCAGAGGAAGCGCTACGACCTGGCCCGCGGTGGCGCTACAGGTCAACAAGAACTGG 60
 DB 1 TCAGAGGAAGCGCTACGACCTGGCCCGCGGTGGCGCTACAGGTCAACAAGAACTGG 60
 OY 61 GCCTGAACCGCGCCGACGATCAGCTGCTGCTACCTGACGAGAGAGAGCTGCGCA 120
 DB 61 GCCTGATGTGCGCGGACGACATCAGCTGCTGACGCTGACGAGAGAAAGAACTGCGCA 120
 OY 121 CCATCGAGTACCTGTGCTGCGCTGACGAGGCGCAGACCATGATGCGCGCGCGCG 180
 DB 121 CCATCGAATATCTGTGCTGCGCTGACGAGGCGTCAACCATGATGCGCTGCGCGCGCG 180
 OY 181 TCGAGGTCCCGGTGAGGTGAGCAGCAT 208
 DB 181 TCGAGGTCCCGGTGAGAAACGACGACAT 208

RESULT 12
 AAT09676
 ID AAT09676 standard; DNA: 970 BP.

AC AAT09676;

DT 15-OCT-1996 (first entry)

DE Mycobacterium tuberculosis rpoB gene DNA sequence.

KM Tuberculosis: disease diagnosis; oligonucleotide; DNA primer; PCR;
 KW polymerase chain reaction; DNA amplification; rpoB locus; TB; ss.

OS Mycobacterium tuberculosis.

FN Key Location/Qualifiers

FT primer_bind 10..27

FT /tag= a

FT /note= "primer FENLFP"

FT /tag= b

FT /note= "primer DDIDL"

FT /tag= c

FT /note= "primer DDIDL"

FT /tag= d

FT /note= "primer rpo95"

FT /tag= e

FT /note= "primer rpo105"

FT /tag= f

FT /note= "primer KY290"

FT /tag= g

FT /note= "M. tuberculosis signature nucleotide"

FT /tag= h

FT misc_feature

FT misc_feature

50	Sequence	3519 BP; 675 A; 1081 C; 1183 G; 580 T; 0 other;
	Query Match	80.0%; Score 166.4; DB 22; Length 3519;
	Best Local Similarity	87.5%; Pred. No. 1.9e-28;
	Matches 182; Conservative	0; Mismatches 26; Indels 0; Gaps
Qy	1	TCAAGGAGAACGCTACGACCTGCGCGCGTGGCGCTACAAAGGTCAACAAAGAAGCTGG 60
Db	884	TCAAGGAGAACGCTACGACCTGCGCGCGTGGCGCTACAAAGGTCAACAAAGAAGCTGG 60
Qy	61	GCGTGAACCGCGGCGCGGCGGATCAGCTGCTGCAGCTCTGACCGAGGAAGACGTCGTCGCCA 120
Db	944	GCGTGCATCTCGGCGAGGCCATCAGCTGCTGCAGCTGACCGAGGAAGACGTCGTCGCCA 120
Qy	121	CCATCGAGTACCTGGTGGCGCTGCACAGAGGCCAGACACGATGACCGTCCCGCGCGCGC 180
Db	1004	CCATCGAATATCTGGTCCGCTTGCACAGAGGTCAGACACGATGACCGTCCCGCGCGCGC 180
Qy	181	TCGAGGTCCCGTGCAGGTGGACGACAT 208
Db	1064	TCGAGGTCCCGTGGAAACGACGACAT 1091
RESULT 14		
AH02079		
ID	AH02079	standard; DNA; 3534 BP.
XX		
AC	AH02079;	
XX		
DT	24-JUL-2001	(first entry)
XX		
DE	Mycobacterium tuberculosis	nucleotide sequence SEQ ID NO:2072.
XX		
KW	Species specific; genus specific; family specific; probe; detection;	
KW	identification; algal; archaeal; bacterial; fungal; parasiticol;	
KW	microorganism; diagnosis; translation elongation factor Tu; toxin;	
KW	translation elongation factor G; RecA recombinase; resistance;	
KW	catalytic subunit of proton-translocating ATPase; antimicrobial;	
KW	vaccine; primer; ds.	
XX		
OS	Mycobacterium tuberculosis.	
XX		
PN	WO200123604-A2.	
XX		
PD	05-APR-2001.	
XX		
PF	28-SEP-2000; 2000WO-CA01150.	
XX		
PR	28-SEP-1999; 99CA-2283458.	
PR	19-MAY-2000; 2000CA-2307010.	
XX		
PA	(INFE-) INFECTIO DIAGNOSTIC (IDI) INC.	
XX		
PI	Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;	
PI	Picard FJ, Roy PH;	
XX		
DR	WPI; 2001-245006/25.	
XX		
XX		
PT	Nucleic acid sequences are used to generate universal probes and	
PT	primers which can be used to identify and detect the presence of algal,	
PT	archaeal, bacterial, fungal and parasiticol species in a test sample -	
XX		
PS	Disclosure; Page 1478-1479; 1580pp; English.	
XX		
CC	The present invention describes a method for generating a repertory of	
CC	nucleic acids of tuf, fus, atpD and/or recA genes from which probes	
CC	and/or primers are derived. The method comprises amplifying the nucleic	
CC	acids of determined algal, archaeal, bacterial, fungal and parasiticol	
CC	species with a combination of defined primer pairs. The method can be	
CC	used for producing probes and/or primers for detecting one or more	
CC	related microorganisms e.g. algae, archaea, bacteria, fungi and	
CC	parasites, for universal detection and for specific and ubiquitous	
CC	detection and identification of an algal, archaeal, bacterial, fungal	

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 ; Search time 636.023 Seconds
(without alignments)
9517.553 Million cell updates/sec

Title: US-09-697-123B-19
Perfect score: 208
Sequence: 1 tcaagagagaagcgctacgac.....ccggctcgagggtgacgacat 208

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1:	gb_ba.*
2:	gb_htg.*
3:	gb_in.*
4:	gb_om.*
5:	gb_ov.*
6:	gb_pat.*
7:	gb_ph.*
8:	gb_pl.*
9:	gb_pr.*
10:	gb_ro.*
11:	gb_sts.*
12:	gb_sy.*
13:	gb_un.*
14:	gb_vi.*
15:	em_ba.*
16:	em_fun.*
17:	em_hum.*
18:	em_in.*
19:	em_mu.*
20:	em_om.*
21:	em_or.*
22:	em_ov.*
23:	em_pat.*
24:	em_ph.*
25:	em_pl.*
26:	em_ro.*
27:	em_sts.*
28:	em_un.*
29:	em_vi.*
30:	em_htg_hum.*
31:	em_htg_inv.*
32:	em_htg_other.*
33:	em_htg_mus.*
34:	em_htg_pln.*
35:	em_htg_rod.*
36:	em_htg_vir.*
37:	em_htg_vrt.*
38:	em_sy.*
39:	em_htgo_hum.*
40:	em_htgo_mus.*
41:	em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	171.2	82.3	3752	1	MSU24494	U24494 Mycobacteri
2	166.4	80.0	610	1	MTU318818	AJ318818 Mycobacte
3	166.4	80.0	610	1	MTU318819	AJ318819 Mycobacte
4	166.4	80.0	616	1	MTU318813	AJ318813 Mycobacte
5	166.4	80.0	618	1	MTU318815	AJ318815 Mycobacte
6	166.4	80.0	618	1	MTU318817	AJ318817 Mycobacte
7	166.4	80.0	633	1	MTU318814	AJ318814 Mycobacte
8	166.4	80.0	637	1	MTU318816	AJ318816 Mycobacte
9	166.4	80.0	639	1	MTU318821	AJ318821 Mycobacte
10	166.4	80.0	970	6	I50706	I50706 Sequence 1
11	166.4	80.0	3534	6	AX111339	AX111339 Sequence
12	166.4	80.0	3853	1	MTU12205	U12205 Mycobacteri
13	166.4	80.0	5084	1	MSGRPOB	L2799 Mycobacteri
14	166.4	80.0	19352	1	AE006964	AE006964 Mycobacte
15	166.4	80.0	19770	1	WTC1376	Z95972 Mycobacteri
16	142.4	68.5	3447	6	AR067447	AR067447 Sequence
17	142.4	68.5	37617	1	MLB1790G	Z14314 M.leprae ge
18	142.4	68.5	348950	1	MLEPRTN7	AL583923 Mycobacte
19	109.2	52.5	3941	1	AF242549	AF242549 Amycolato
20	88	42.3	3495	6	AX120631	AX120631 Sequence
21	88	42.3	328050	1	AP005275	AP005275 Corynebac
22	88	42.3	349980	6	AX127144	AX127144 Sequence
23	86.8	41.7	32923	1	SCDB2	AL160431 Streptomy
24	51.2	24.6	204050	1	AL646073	AL646073 Ralstonia
25	48.2	23.2	204050	1	AL646070	AL646070 Ralstonia
26	47.4	22.8	32748	1	AB070951	AB070951 Streptomy
27	47.2	22.7	7954	1	TAQ19223	Y19223 Thermus aqu
28	46.6	22.4	1416	1	SRRECAGEN	X94233 S.rimosus r
29	46.2	22.2	20054	1	SCCB12	AL391588 Streptomy
30	45.8	22.0	329100	1	SME591787	AL591787 Sinorhizo
31	45.6	21.9	1271	1	SVU04837	U04837 Streptomyce
32	45	21.6	134816	2	CNS08C9R	AL772413 Oryza sat
33	45	21.6	135378	2	CNS07YP9	AL713941 Oryza sat
34	44.8	21.5	1952	1	SCU42212	U42212 Streptomyce
35	44.4	21.3	1782	14	AF171937	AF171937 Suid herp
36	44.2	21.3	1734	6	I04901	I04901 Sequence 6
37	44.2	21.3	12112	1	AE005722	AE005722 Caulobact
38	44.2	21.3	92509	1	AL646086	AL646086 Ralstonia
39	44.2	21.3	346294	1	AP002999	AP002999 Mesorhizo
40	44	21.2	1490	1	SARECAG	Z30324 S.ambofacie
41	44	21.2	3488	1	SCWHIB	X62287 S.coelicolo
42	44	21.2	37898	1	SCE34	AL353862 Streptomy
43	43.8	21.1	1440	8	AY034826	AY034826 Curvulari
44	43.8	21.1	2206	3	AY102669	AY102669 Drosophil
45	43.8	21.1	9302	1	SGR243674	AJ243674 Streptomy

ALIGNMENTS

RESULT 1	MSU24494	3752 bp	DNA	linear	BCT 02-MAR-2000
LOCUS	Mycobacterium smegmatis	DNA polymerase (rpoB) gene, complete cds.			
DEFINITION	U24494				
ACCESSION	U24494.1	GI:790347			
VERSION					
KEYWORDS					
SOURCE	Mycobacterium smegmatis.				
ORGANISM	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;				
	Actinomycetales; Corynebacterineae; Mycobacteriaceae;				
	Mycobacterium.				
REFERENCE	1 (bases 1 to 3752)				
AUTHORS	Hetherington,S.V., Watson,A.S. and Patrick,C.C.				
TITLE	Sequence and analysis of the rpoB gene of Mycobacterium smegmatis				

JOURNAL Antimicrob. Agents Chemother. 39 (9), 2164-2166 (1995)
 MEDLINE 96050766
 PUBMED 8540740
 REFERENCE 2 (bases 1 to 3752)
 AUTHORS Hetherington, S.V.
 TITLE Direct Submission
 JOURNAL Submitted (11-APR-1995) Seth V. Hetherington, Infectious Diseases,
 St. Jude Children's Research Hospital, 332 N. Lauderdale, Memphis,
 TN 38101, USA

FEATURES

source Location/Qualifiers
 1..3752
 /organism="Mycobacterium smegmatis"
 /db_xref="taxon:1772"
 194..3703
 /gene="rpoB"
 194..3703
 /gene="rpoB"
 /codon_start=1
 /evidence=experimental
 /transl_table=11
 /product="DNA polymerase"
 /protein_id="AAA91426.2"
 /db_xref="GI:7144498"
 /translation="MLEGCILAVSSQSKNAITNNSVRGAPNGVSFAKLRPLEVRGL
 LDVQTSFEMLVGSRVRAQAIIDRGEENPVGLLEAEVLAELSPIDFSGMSLSFSDPR
 FDEVKASDECKDKMTYAAPLFVTAETFINNNTGEIKSOTVFMGDFPMNTEKGTIIN
 GTERVVSQLVSPGVYFDETDIKSTEKTLHSVKVIPGRGANLEFDVDKRDYGVRIID
 RKRQPVVLLKALGTNEQIVVERGFSEIMGTLEKDTSTGTDALLDIYKRLRPE
 PPTKESAOTLLENLFKEKRYDLARVGRYKVKKLGNAGKPTISTLTPEEDVATIE
 YLVRLEGQTSMTVPFGGVFEFVASHDIDHFRNRRTVGTGELIQNIRVLSRMRVVR
 ERMTQDVAITPTQLINIRPVVAAIKFEFGTSQSPMDQNNPLSGLTHKRLSALG
 PGGLSERRAGLEVDPVHPHYGMCPIETPEGNIGLISLSVYARVNPFGFIETPYR
 KVVGVTDQIDYLTADDEDRHVAQANSPTDENGRTEDRVNVRKKGVEFEVSADO
 VYMDVSPROMYSVATAMIPFLEHDDANALMGANNOQAVPLVSEAPLVGTGHELR
 AATDVTWSSHKTVTEEVSADYITVADDGTQSYRLKFKFARSHNCTANQRPVDA
 GORVEAGOVADGPTQNGEMALGNLLVAIMPWEHNTEDAIILNSRLVEDLTSI
 HHEEIDARDKLAEEITRVPKLSDEVLADDEPRIVRIGAEVRDGEILVGVKTP
 KGTEUTPEERLLRATFGEKAREVRDTSUKVPHGESGVIGIRVFSREDDDELPGVYN
 ELVRVYVQAKRISDGDKLARNHNGKVGITKILPVDEMPFLPDGTPVDIILNTHGCRV
 VLNIGQILTHLGWAKAGNIDVLVAGPDWASKLPEELYSAPADSTVATPVFDCAOE
 GELAGLLGSTLPNRDGEVMVNDGRATLPDGRSSGEPPFPYPTVGYMYILKLHLHVD
 KIHARSTGYSMTIOPLGKQAGFGGQRFGEWECMAQYGAAYTLOELLTKISDDTV
 GRVKYEAIVKGENIPGCPESFEKVLKELQSLNLNVEVSSDGRAEMRDGDDDL
 ERAANLGINLSRESASVEDLA"

BASE COUNT 706 a 1225 c 1210 g 611 t

ORIGIN

Query Match 82.3%; Score 171.2; DB 1; Length 3752;
 Best Local Similarity 88.9%; Pred. No. 4, 6e-19;
 Matches 185; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTGGCCGCTACAGGTCAACAAAGAGCTGG 60
 DB 1068 TCAAGGAGAGCGCTACGACCTGGCCCGGTGGCCGCTACAGGTCAACAAAGAGCTGG 1127
 QY 61 GCCTGAACGCGCGCAGCGCATCAGCTGCTGCGACTCTGACCCGAGGAGACGCTGTCGCCA 120
 DB 1128 TCGTGAACGCGCGCAGCGCATCAGCTGCTGCGACTCTGACCCGAGGAGACGCTGTCGCCA 1187
 QY 121 CCATCGAGTACCTGGTGGCGCTCAGCAGGCGGCACACGACGATGACCTCCCGCGCGCGG 180
 DB 1188 CCATCGAGTACCTGGTGGCGCTCAGCAGGCGGCACACGACGATGACCTCCCGCGCGCGG 1247
 QY 181 TCGAGGTCCCGGTGCGAGGTGGAGGACAT 208
 DB 1248 TCGAGGTCCCGGTGCGAGGTGGAGGACAT 1275

RESULT 2

MTU318818 610 bp DNA circular BCT 09-AUG-2002
 LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
 DEFINITION beta subunit, isolate 1415-97.

ACCESSION

AJ318818
 AJ318818.1 GI:22208412
 RNA polymerase beta subunit; rpoB gene.
 Mycobacterium tuberculosis.
 Mycobacterium tuberculosis.
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
 tuberculosis complex.

REFERENCE

AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
 TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
 isolated in Spain (1996-2001). Description of new alleles into rpoB
 gene and review
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 610)
 AUTHORS Herrera, L.
 TITLE Direct Submission
 JOURNAL Submitted (07-AUG-2001) Herrera, L., Servicio Bacteriologia, Centro
 Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
 Majadahonda. Madrid. 28220, SPAIN

FEATURES

source Location/Qualifiers
 1..610
 /organism="Mycobacterium tuberculosis"
 /isolate="1415-97"
 /db_xref="taxon:1773"
 1..610
 /gene="rpoB"
 <1..>610
 /gene="rpoB"
 /codon_start=1
 /transl_table=11
 /product="RNA polymerase beta subunit"
 /protein_id="CAC87035.1"
 /db_xref="GI:22208413"
 /translation="LDIVYRLRPGEPPTKESAOITLLENLFFKRYDLARVGRYKVNK
 KLGRVGEPIQTSSTLTEDVVATLEYLVRLHEGTTMTVPGGVPEVPTDDIDHFGNR
 RLRTVGEPIQNRVGRMRVVRVRRMTQDVAITPTQLINIRPVVAAIKFEFGTS
 OPSQFMGONPLSLTHKRLSALGPGGLSRRERAGLEVDRVHP"

BASE COUNT

122 a 191 c 202 g 95 t

Query Match

Best Local Similarity 80.0%; Score 166.4; DB 1; Length 610;
 Best Local Similarity 87.5%; Pred. No. 3, 8e-18;
 Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTGGCCGCTACAGGTCAACAAAGAGCTGG 60
 DB 80 TCAAGGAGAGCGCTACGACCTGGCCCGGTGGCCGCTACAGGTCAACAAAGAGCTGG 139
 QY 61 GCCTGAACGCGCGCAGCGCATCAGCTGCTGCGACTCTGACCCGAGGAGACGCTGTCGCCA 120
 DB 140 GGCTGCATGTGCGGAGGCCCATCAGCTGCTGCGACTGACCCGAGGAGACGCTGTCGCCA 199
 QY 121 CCATCGAGTACCTGGTGGCGCTCAGCAGGCGGCACACGATGACCTCCCGCGCGCGG 180
 DB 200 CCATCGAATATCTGCTCGCTTGCACGAGGCTCAGACCATGACCTGTCGCGCGCGG 259
 QY 181 TCGAGGTCCCGGTGCGAGGTGGAGGACAT 208
 DB 260 TCGAGGTCCCGGTGGAACCGAGGACAT 287

RESULT 3

MTU318819 610 bp DNA circular BCT 09-AUG-2002
 LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
 DEFINITION beta subunit, isolate 1417-97.
 ACCESSION AJ318819
 VERSION AJ318819.1 GI:22208414
 KEYWORDS RNA polymerase beta subunit; rpoB gene.
 SOURCE Mycobacterium tuberculosis.
 ORGANISM Mycobacterium tuberculosis.
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;


```
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid: 28220, SPAIN
Location/Qualifiers
1..618
/organism="Mycobacterium tuberculosis"
/isolate="2540-97"
/db_xref="taxon:1773"
1..618
/gene="rpoB"
<1..>618
/codon_start=1
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"
/translation="TDEALLDIYKLRPGEPPTKESAOITLLENLFFKRYDLARVGR
YKNNKLGHLHVGEPITSTLTEDDVVATIEYLRLHGGQTTMTVPGGVEVPVETDDID
HFGNRLRTYVGLIQIRVGSRMERVRMTTQDVEAITPTRLINIRPVAAIKE
FFGTSQLSQFMDQNNPLSLGTHKRRLALPGGLSRERAGLEVRDV"
BASE COUNT 124 a 190 c 207 g 97 t
ORIGIN
Query Match 80.0%; Score 166.4; DB 1; Length 618;
Best Local Similarity 87.5%; Pred. No. 3.8e-18;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 TCAAGGAGAACGCTACGACCTGGCCGCGTGGCGCTACAGGTCAACAAGAGCTGG 60
|||||
DB 95 TCAAGGAGAACGCTACGACCTGGCCGCGTGGCGCTACAGGTCAACAAGAGCTGG 154
|||||
QY 61 GCCTGAACGCCGCCAGCCGATCAGCTCGTCTGACCTCGACCGAGGAGAGCTCGTCCGCA 120
|||||
DB 155 GCCTGATGTCGGCGAGCCATCAGCTCGTCTGACCGAGGAGAGAGCTCGTCCGCA 214
|||||
QY 121 CCATCGAGTACCTGTCGCTGCACGAGGCGCAGACGATGACCGTCCCGCGCGCG 180
|||||
DB 215 CCATCGAATATCTGTCGCTGACGAGGTCAGACGATGACCGTCCCGCGCGCG 274
|||||
QY 181 TCGAGTCCCGTTCGAGGTGGAGACAT 208
|||||
DB 275 TCGAGTCCCGTGGAAACGAGACAT 302
|||||
RESULT 6
MTU318817 618 bp DNA circular BCT 09-AUG-2002
LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
DEFINITION beta subunit, isolate 2348-98.
ACCESSION AJ318817
VERSION AJ318817.1 GI:22208410
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1
REFERENCE
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL Unpublished
REFERENCE
AUTHORS Herrera, L.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid: 28220, SPAIN
Location/Qualifiers
1..618
/organism="Mycobacterium tuberculosis"
/isolate="2348-98"
FEATURES
source
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid: 28220, SPAIN
Location/Qualifiers
1..618
/organism="Mycobacterium tuberculosis"
/isolate="2540-97"
/db_xref="taxon:1773"
1..618
/gene="rpoB"
<1..>618
/codon_start=1
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"
/translation="TDEALLDIYKLRPGEPPTKESAOITLLENLFFKRYDLARVGR
YKNNKLGHLHVGEPITSTLTEDDVVATIEYLRLHGGQTTMTVPGGVEVPVETDDID
HFGNRLRTYVGLIQIRVGSRMERVRMTTQDVEAITPTRLINIRPVAAIKE
FFGTSQLSQFMDQNNPLSLGTHKRRLALPGGLSRERAGLEVRDV"
BASE COUNT 124 a 190 c 207 g 97 t
ORIGIN
Query Match 80.0%; Score 166.4; DB 1; Length 618;
Best Local Similarity 87.5%; Pred. No. 3.8e-18;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 TCAAGGAGAACGCTACGACCTGGCCGCGTGGCGCTACAGGTCAACAAGAGCTGG 60
|||||
DB 95 TCAAGGAGAACGCTACGACCTGGCCGCGTGGCGCTACAGGTCAACAAGAGCTGG 154
|||||
QY 61 GCCTGAACGCCGCCAGCCGATCAGCTCGTCTGACCTCGACCGAGGAGAGCTCGTCCGCA 120
|||||
DB 155 GCCTGATGTCGGCGAGCCATCAGCTCGTCTGACCGAGGAGAGAGCTCGTCCGCA 214
|||||
QY 121 CCATCGAGTACCTGTCGCTGCACGAGGCGCAGACGATGACCGTCCCGCGCGCG 180
|||||
DB 215 CCATCGAATATCTGTCGCTGACGAGGTCAGACGATGACCGTCCCGCGCGCG 274
|||||
QY 181 TCGAGTCCCGTTCGAGGTGGAGACAT 208
|||||
DB 275 TCGAGTCCCGTGGAAACGAGACAT 302
|||||
RESULT 7
MTU318814 633 bp DNA circular BCT 09-AUG-2002
LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
DEFINITION beta subunit, isolate 1058-97.
ACCESSION AJ318814
VERSION AJ318814.1 GI:22208404
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1
REFERENCE
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL Unpublished
REFERENCE
AUTHORS Herrera, L.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid: 28220, SPAIN
Location/Qualifiers
1..633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1..633
/gene="rpoB"
<1..>633
/codon_start=1
FEATURES
source
```


/product="RNA-polymerase beta subunit"
/protein_id="AA020242.2"
/db_xref="GI:714499"
/translation="MLEGCIADSRQSKTAASPSRQSSNNNSVPGAPNRVSAFLK
REPLVPGLLDVOQDSFVWLGISPMRESAAERGDVNPVGGLEEVLYELSPEDFSGS
MSLSFSDPRDDYKAPVDECKDKMTYAAPLFVTAEFINNNTGEIKSOTVFMGDFPM
TEKGTFLINGTERVVSQVSPGVFDEITDKSTDKTLHSVKVIPSRGAWLEFVDK
RDVTVGRVDRKRQPVTVLLKALGWTSEQIVERFGFSEIMSRSTLEKNDTGTDEALLD
IYKRLRGPDPPTKSAQTLLENLFFKFKRYDLARVGRYKVNKLGHVGEPTTSLLT
EEDVATIEVLRHLEGOTTMTVPGGVEPVETDDIDHFGNRLKRLTVGELIQIRVG
MSMRVVRERMTTQDVEAITPOTLINIRPVAAIKFEFCTSQLSOFMQNPLSGLT
HKRLSALGGLRERAGLEVHPSHYHRCMPTETPEGNIGLISLVSARVNP
FGYETPRKVDGVSDGIVLTADEEDRHVVAOANSPIADGRVPRVLVRRKAG
EVGVPSSEVDYMDVSPROMSVATAMIPFLEHDDANRALMGANROQAVPLVSEAP
LVGTGMLRAAIDAATSSQESGVIIEVSADYITVMHDNGTRTYRMKRFARSNGTC
ANOCPIVDAGDRVEAGQVADGCTDDGEMALGNLLVAIMPWEGHNYEDAIIILSNRL
VEEDVLTSHIEHEIDARDTKLGAEEITRDIPTNISDEVLAIDLDERGIVRIGAEVRDG
DILVGVTPKGETELTPEERLLRAIFGEKAREVRDTSKVPHGSGKVIIGIRFVSRED
EDELPAVGNELRVVYVAQRKTSDDGKLAGRHNGKVIIGKILPVEDMPFADGTVDI
ILNTHGVPRRMNTQIETHLGCWASHGKVDAAKGVDPWAARLPDELLEAOPNAIVS
TPVFDGAQAEALQGLLSCTLPNRDGLVDADGKAMLPDGRSEPEPPVPTVGYMYIM
KLHLVDDKIHARSTGYSMTIQOPLGGRKQAGGQRFGEWECWAMQAYGAATYLOELL
TIKS"

BASE COUNT 723 a 1173 c 1293 g 664 t
ORIGIN
Query Match 80.0%; Score 166.4; DB 1; Length 3853;
Best Local Similarity 87.5%; Pred. No. 2.8e-18;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAGCGGTACGACCTGGCCCGCTGAGCGGCTTACAAAGTCAACAAGAGCTGG 60
Db 1477 TCAAGGAGAGCGGTACGACCTGGCCCGCTGCTTAAAGTCAACAAGAGCTCG 1536
Qy 61 GCCTGAACGGCGGCGACCATACGTCGACTCTGACCGAGGAGAGCTGCTGCCCA 120
Db 1537 GCCTGCATCTGCGGACCCATCAGTCGTCGACGCTGACCGAAGAGAGCTGCTGCCCA 1596
Qy 121 CCATCGAGTACCTGGTGGCTGCACGAGGCGGCGACGATGACCGTCCCGCGGCGCG 180
Db 1597 CCATCGAATATCTGTCGCTTGCACGAGGCTCAGACGATGACCGTTCGCGGCGCG 1656
Qy 181 TCGAGGTCGCGGTCGAGTGAGACAT 208
Db 1657 TCGAGGTGCGCGTGGAAACCGAGACAT 1684

RESULT 13
MSGRPOB 5084 bp DNA linear BCT 13-SEP-1994
LOCUS
DEFINITION
Mycobacterium tuberculosis RNA polymerase beta-subunit (rpoB)
gene, complete cds and RNA polymerase beta'-subunit rpoC gene,
partial cds.
ACCESSION L27989.1 GI:468333
VERSION
KEYWORDS
SOURCE
ORGANISM
Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1 (bases 1 to 5084)
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
FEATURES
Source
1..5084
/organism="Mycobacterium tuberculosis"
/strain="Rv"
/db_xref="taxon:1773"
1065..4598

CDS

/gene="rpoB"
1065..4598
/gene="rpoB"
/codon_start=1
/evidence-experimental
/transl_table=11
/product="RNA polymerase beta-subunit"
/protein_id="AA021416.1"
/db_xref="GI:468334"
/translation="MLEGCIADSRQSKTAASPSRQSSNNNSVPGAPNRVSAFLK
REPLVPGLLDVOQDSFVWLGISPMRESAAERGDVNPVGGLEEVLYELSPEDFSGS
MSLSFSDPRDDYKAPVDECKDKMTYAAPLFVTAEFINNNTGEIKSOTVFMGDFPM
TEKGTFLINGTERVVSQVSPGVFDEITDKSTDKTLHSVKVIPSRGAWLEFVDK
RDVTVGRVDRKRQPVTVLLKALGWTSEQIVERFGFSEIMSRSTLEKNDTGTDEALLD
IYKRLRGPDPPTKSAQTLLENLFFKFKRYDLARVGRYKVNKLGHVGEPTTSLLT
EEDVATIEVLRHLEGOTTMTVPGGVEPVETDDIDHFGNRLKRLTVGELIQIRVG
MSMRVVRERMTTQDVEAITPOTLINIRPVAAIKFEFCTSQLSOFMQNPLSGLT
HKRLSALGGLRERAGLEVHPSHYHRCMPTETPEGNIGLISLVSARVNP
FGYETPRKVDGVSDGIVLTADEEDRHVVAOANSPIADGRVPRVLVRRKAG
EVGVPSSEVDYMDVSPROMSVATAMIPFLEHDDANRALMGANROQAVPLVSEAP
LVGTGMLRAAIDAATSSQESGVIIEVSADYITVMHDNGTRTYRMKRFARSNGTC
ANOCPIVDAGDRVEAGQVADGCTDDGEMALGNLLVAIMPWEGHNYEDAIIILSNRL
VEEDVLTSHIEHEIDARDTKLGAEEITRDIPTNISDEVLAIDLDERGIVRIGAEVRDG
DILVGVTPKGETELTPEERLLRAIFGEKAREVRDTSKVPHGSGKVIIGIRFVSRED
EDELPAVGNELRVVYVAQRKTSDDGKLAGRHNGKVIIGKILPVEDMPFADGTVDI
ILNTHGVPRRMNTQIETHLGCWASHGKVDAAKGVDPWAARLPDELLEAOPNAIVS
TPVFDGAQAEALQGLLSCTLPNRDGLVDADGKAMLPDGRSEPEPPVPTVGYMYIM
KLHLVDDKIHARSTGYSMTIQOPLGGRKQAGGQRFGEWECWAMQAYGAATYLOELL
TIKSDDGLERAANLGINLSNESASPEDLA"
4641..5084
/gene="rpoC"
4641..>5084
/gene="rpoC"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta'-subunit"
/protein_id="AA021417.1"
/db_xref="GI:537608"
/translation="MLDVNPFDELRLIGLATAEDIRWISYGEVKKPVTINVTTLKPEKD
GLFCEKIFGPTRDWECYCGYKRVRFKGIICERCGVEVTRAKVRERBMGHIELAAPVT
HIWYFKGVPSRLGYLLDLAPKDLKIIYFAAYVITSVDEEMRHNEL"
BASE COUNT 969 a 1534 c 1691 g 890 t
ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 5084;
Best Local Similarity 87.5%; Pred. No. 2.7e-18;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAGCGGTACGACCTGGCCCGCTGAGCGGCTTACAAAGTCAACAAGAGCTGG 60
Db 1966 TCAAGGAGAGCGGTACGACCTGGCCCGCTGAGCGGCTTACAAAGTCAACAAGAGCTGG 2025
Qy 61 GCCTGAACGGCGGCGACCATCAGTCGCTGCGACTCTGACCGAGGAGACGCTGCTGCCCA 120
Db 2026 GCCTGCATCTGCGGAGCGCCATCAGTCGTCGACGCTGACCGAAGAGACGCTGCTGCCCA 2085
Qy 121 CCATCGACTACTGCTGCGCTGTCAGGAGCGCCAGACCATGACCGTCCCGCGCGCGC 180
Db 2086 CCATCGAATATCTGCTGCGCTTGCACGAGGCTCAGACCATGACCGTTCGCGGCGCGC 2145
Qy 181 TCGAGGTCCCGCTGAGGTGGACGACAT 208
Db 2146 TCGAGGTCCCGTGGAAACCGACGACAT 2173

RESULT 14
AE006964
LOCUS
DEFINITION
Mycobacterium tuberculosis CDC1551, section 50 of 280 of the
complete genome.
ACCESSION
VERSION
AE006964 AE000516
AE006964.1 GI:13880217

KEYWORDS	Mycobacterium tuberculosis CDC1551.
SOURCE	Mycobacterium tuberculosis CDC1551
ORGANISM	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE	1 (bases 1 to 19352)
AUTHORS	Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
TITLE	Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 19352)
AUTHORS	Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
TITLE	Direct Submission
JOURNAL	Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES	Location/Qualifiers
source	1. 19352 /organism="Mycobacterium tuberculosis CDC1551" /strain="CDC1551" /db_xref="taxon:83331" /note="clinical strain" 163..3699 /gene="MT0695" 163..3699 /gene="MT0695" /note="similar to GB:L27989 GB:L05910 GB:U12205 SP:P47766 PID:149992; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="DNA-directed RNA polymerase, beta subunit" /protein_id="AAK44921.1" /db_xref="GI:13880218"
gene	/translation="MLEGCCILADSRQSKTAASPSRQSSNNSVPGAPNRVSFAKL REPLEVGLDQVTDSEFWLIGSPRMSAAERGDNVPVGLLEEVLYELSEIDPFSGS MSLSFSDPDDKAPVDECKDKMTYAAPLFVTAEFINNNTGKTSQVFMGDFPMX TEKGTFLINTERVVYQSLVSPGVYFDETIDKSTDKTLHSVKVIPSGANLEEDVDK RDYGVRIIDPKRPQVTVLLKALGWTSQIVERRGFSEIMRSTLEKONTVGTDEALLD IYKRLGPEPTKESAQTLLENLFFKAKYDILARVRYKVNKKLGLHVGEPTTSITLT EDVYATIEVLRHGGQTTMTVPGGVEPVETDDIDHGNRRLRTVGELIQNQIRVG MSRMRVYRERTTQDVEAITPQTLINIPVYVAAIKEFFGTSQLSQFMDQNNPLSGLT HKRRLSALGPGSLRERAGLEVDPVSHYGRMCPTEPEGNIGLSLSVYARVNP EVEYVPSSEVDYMDVSPROMSVATAMIPFLEHDDANRALKGANNQORQAVPLVRSEAP LVCTGMELRAADAGDVVVAESGVIEEVSADYITVMDHNGTRRTYRMKFKARNHGT CANQCIFVADGRVEAGQVIADGCPCTDDGEMALGNLLIIVMPWEGHNYEDAIIISNR LVSEEDVLTSHIEEHEIDARDTKLGAEEITRDIPNISDEVLADLDREIVRIGAEVRD GDLGVKPTPKGETVLTPEERLLRAIFGKAREVRDTSKVPHGSGKVGIRVFSRE DEDELPAQVNLVYVVAQKRIISDQGLAGRHNGKVGIGILPVEDMPTLADCTPVD IILNTHGVPRRNIGQILETHLWGCAHSCWKVDAAKGVPDMAARLPDLLEAQNAIV STPVFDGAQAEQGLISCTLPNROGDVLVDAGKAMLFDRSGSEPPYPVYVTGIMLI MKLHLVDRKLIHARSTGPPSHMTQPLGKAGQGGQREGMECWAMQAYGAAYTLMEL LTKSDDTGVRVYEAIVKGENIPEGPISFESFVLLKELQSLCLNVEVLSDDGAATE LRGEDELRAAANLGINLSRNASVEDLA"
gene	3744..7694 /gene="MT0696"
CDS	3744..7694 /gene="MT0696" /note="similar to SP:P37871; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="DNA-directed RNA polymerase, beta-prime subunit" /protein_id="AAK44922.1" /db_xref="GI:13880219"
gene	/translation="MLDVNFDFDELRIGLATAEDIRQMSYGEVKKPKETINRYTLKPKED GLFCEKIFGPTDMECYCGKYKVRREKGLICERCYGEVETRAKVRERMGHIELAAPVT HINYFKGVP SRLCYLLDAPKDEKLIYFAAYVITSVDEMRHNEUJLEAEVAVERK AVEDORGELEAQAQLEADLAELAEAGAKARRKVRDGGERMRQINDRAOERLDR LEDWIRFTKAPLQKQILVDENLYRELVDYGEYFTTGAMGAEISIOKLIENFDIAEAS LDRVIRNGKQKKLRALKRLKVVAAPQSGNSPMGVMDAVPVIPELQPMVQLDQGR FATSDNLRYRVINNRNRLKRLIDLCAPETIVNNEKRLQESVDALFDNGRGRPV PGNRP LKSLD LKQGRFRONLIGKRDVSGRSVIVVGPQLKHOGGLPKMALE LQAFPMVKRLVDNLNHAQNLKSARKYERQRPQVMDVLEEVIAEHPVLNARPTLHRLG IOAFPEMLVGEKAIQHLPLCEAFNADFDQDNVHPLPLSAQAQAEARILMSSNNIL SPASRPLAMPRLDMVTGLYLTVEPDETYEQPASGDHPETGYVSPAEAIMAADR FVLSVRAKIKVRLTQLPVPEIAELFSGHSGQPGDMAETTLGRVMFNELLPLGYP VKNQHKHKYQAAIINDLAERYPMIVAVOTVDKLDAGFYATRCGVTSVSMADVLVPP RKKEILDHYEADNKVEKOFORGALNHDRNEALVEIKKEATDEVGQALREHYPDNP IITIVDSGATGNFTQTRTAGMKGLVTPKGEIIRPPKSSPREGTLVLEFINTHGA RKLADTAFATDSGYLTRLDVDSQDVIIVREHDCQTERGIVVLEAERAPDGTILROP YIETSAFTALGTADVADEAGNVIVERGQDLQDEIDALLAAGITQVKVRSVLTCATST GICATCYGRSMATGKLVLDIGEAIVGAQSIGEPQDITLMRTFHOGGVGEDITGGLPR VQELFARVPRGKAPADYTVGRVRLDEGRFYKITVPDDGGEVYDOKISKRQRLRV FKHEDSERVLSGDHVEGQQLMEGSADPHVLRVQGPREVQIHLVREVEVYRAQG VSTHDKHIEVIRQMLRRVTIIDSGSTFLPGSLIDRAEFEEAENRVRVVEGGEPAAGR PVLMGITKIASLATDSWLSAASFQETTRVLTDAAINCRSDKLNGLKENVIKGLIPAGT GINRYRNAVQPTPEERAAAYTIPSTYEDQYISPDFGAATGAAPVLUDDYGYSDYR" complement(7691..8065) /gene="MT0697" complement(7691..8065) /gene="MT0697" /note="identified by Glimmer2; putative" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAK44923.1" /db_xref="GI:13880220"
CDS	/translation="MFDAAALNPGHAWASAMERSGLLECAGLDEQPFGEFTADKL NPDGSSRRVPRRQADGGTATVHERGGQSGGAGVVPQMHGFPALAMQDRLIHH GEOTQNRIAQAFNRVFCVCSPT" complement(8058..9972) /gene="MT0698" /note="This region contains an authentic frame shift and is not the result of a sequencing artifact; identified by Glimmer2; putative; conserved hypothetical protein, authentic frameshift" 10167..10925 /gene="MT0699" 10167..10925 /gene="MT0699" /note="identified by match to PFAM protein family HMM pf01261" /codon_start=1 /transl_table=11 /product="AP endonuclease, family 2" /protein_id="AAK44924.1" /db_xref="GI:13880221"
gene	/translation="MLGSHVSPDPLAAAEAGADVVOIFLGNPQSWKAPKPRDDAA AKKATLPIYVHAPYLINLASANNRIPSRKILQETCAAAADIGAAAVIVHGGHVD DNDIDKGFQNRKALDTEVPEVYLENTAGDAMARRFDITIALMDVIGDTGIGFC LDTCHTAAAGEALTDAVRIDRKAITRIDLVHCNDSRSEAGSRDRHANLAGSGQIDPDL LVAAVKAAGAPVICETADQGRKDDIAFLRERTGS" 10957..11799 /gene="MT0700" 10957..11799 /gene="MT0700" /note="similar to GB:U00012 PID:466863; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="hydrolase/esterase, putative" /protein_id="AAK44925.1" /db_xref="GI:13880222"
gene	/translation="MLRRVAILLAAVLAFAAGCSGGTRLAAGFGNGNSVHTLDVGDAGR SYRLKXPGVGLSPSAPVYMLHGGFGSAQAERSYGMDELADSEKFLVAYPDGVHRAWN ANGGCGCCGRPAREGVDDIGFVRAVVADIANNVSDIPARVYVTGNSGAINSYTLACWT SIFAALGVVSGTQDPCQSPRPVSVIHHTGTADPLVRYHGGPCAGFARIDGPPVDUN

AFWREVNRCALDITTEGPVTTSGATCADNRRVLLTVDDAGHRWPSFATQTLWRFFA
AHRF"

gene 11859. .13487

CDS /gene="MT0701"

11859. .13487

/gene="MT0701"

/note="similar to SP:P33224 GB:L20915 PID:457172

PID:457174 PID:537028; identified by sequence similarity;

putative"

/codon_start=1

/transl_table=11

/product="acyl-CoA dehydrogenase, putative"

/protein_id="AAK44926.1"

/db_xref="GI:13880223"

/translation="MSDTHVTNQVPPLENYNPASSPVLEALIOEGQWGLDEVNEV

GAISACQQRWQELADNRNRPILRTHDGYRVDEYDPAYHELMRTAITHGMAAP

WADRGCAHVRAAKTSVMTVEPHICPISMTVAVVPALRYNSLAAYVEPLLTSPREY

DEPKPATKAGITAGMSMTKEKOGSDVRAGTQATPNADGSYSLTGKWFISAPMCD

IFVLQAQPDGLSCFLLPVLPDGNRMFLQRLKDKLGNHANASSEVEYDGAVALV

GEGRGVPIIENVNLTLDCAIGSATSMRTGLTRAVHQAHRKAFGAYLIDPOLMRN

VLADLAVERAETIVAMRMAGATDNVRNTEALLRRIGLAAAKYVWCKRSTAHAAE

ALBCLNGYVEDSGPRPLYREAPLMGIWEGSNVSALDTRAMATRPACVEVLDEL

ARSAGQDPRLDGHVERLRPLQGLDITIGYRARKIAEDICLALQGLSLVRRHGPAAVEA

FLATRLGGWGGAYGTMFAGLDLAPILERALYKG"

13498. .14436

/gene="MT0702"

13498. .14436

/gene="MT0702"

/note="similar to GP:3885480; identified by sequence

similarity; putative"

/codon_start=1

/transl_table=11

/product="enoyl-CoA hydratase/isomerase family protein"

/protein_id="AAK44927.1"

/db_xref="GI:13880224"

/translation="MTHAIRPVDNLKMTYEVTRGRIATFNRPKGNATIAIDTPL

ELSAVERADLPQGVHIVILVSRGEGFCFADLSAYAGSSSTGGGAYQGVLDGKT

QAVNHLPNQWPMIDYOMKSRFVRGFCASLMHADRPVYKINGCYVAGGDTIALRADQ

VIAAADAKGYPTPRVWGPVPAAGLWHRILGQDQAKRLFLTGDCITGAQAAGLAVEA

PEPADLTERLVAIALPQLIMVKLAINLSALLOOGVATSRMVSIVFDGAARHT

PEHAFAVADAVEHGFDAVRRDEPFGDYGRQASRV"

14439. .15161

/gene="MT0703"

14439. .15161

/gene="MT0703"

/note="identified by Glimmer2; putative"

/codon_start=1

/transl_table=11

/product="hypothetical protein"

Query Match 80.0%; Score 166.4; DB 1; Length 19352;

Best Local Similarity 87.5%; Pred. No. 2.2e-18;

Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTGCCCGCGCTGGCGCGCTACAGGTCAACAAGAGCTGG 60

Db 1064 TCAAGGAGAGCGCTAGGACCTGCCCGCGCTGGCTATAGGTCAACAAGAGCTCG 1123

QY 61 GCCTGAACGCCGCCAGCCGATCAGCTGCTGCTACTCGACCGAGGAGACGCTCGTGC 120

Db 1124 GGCTGCTATGTCGGCGAGCCCATCACCTGCTGCTGCTGCTGCTGCTGCTGCTG 1183

QY 121 CCATCGAGTACCTGGTGGCGCTCAGGAGGCCAGACCGATGACCGTCCCGGGGGGG 180

Db 1184 CCATCGAATATCTGGTGGCGCTTCCAGAGGGTCAGACCGATGACCGTCCCGGGGGGG 1243

QY 181 TCGAGGTCCCGGTCGAGGTCGACGACAT 208

Db 1244 TCGAGGTCCCGGTGGAACCGACGACAT 1271

RESULT 15
MTC1376

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

RBS

gene

CDS

MTC1376
Mycobacterium tuberculosis H37Rv complete genome; segment 32/162.
295972 AL123456
295972.1 GI:3261790

Mycobacterium tuberculosis H37Rv.
Mycobacterium tuberculosis H37Rv
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium tuberculosis complex.

1 (bases 1 to 19770)
Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,
Harris, D., Gordon, S.V., Eiglmeyer, K., Gas, S., Barry, III, C.E.,
Tekala, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,
Connor, R., Davies, R., Devlin, K., Felwell, T., Gentles, S.,
Hamlin, N., Holroyd, S., Hornsby, T., Jagsels, K., Krogh, A., McLean, J.,
Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A.,
Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, S.,
Squares, S., Squares, R., Sulston, J.E., Taylor, K., Whitehead, S. and
Barrell, B.G.

Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence
Nature 393 (6685), 537-544 (1998)

98295987
9634230
2 (bases 1 to 19770)
Parkhill, J.

Direct Submission
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 27, 1998 this sequence version replaced gi:2143285.

Notes:
Details of M. tuberculosis sequencing at the Sanger Centre are
available on the World Wide Web.

(URL, http://www.sanger.ac.uk/projects/M_tuberculosis/) CDS have
been renumbered from the original cosmid submissions but the old
gene designations are in brackets after the new gene numbers.
Gene prediction was based on a Hidden Markov Model of TB genes
implemented in TBparse (Krogh) supplemented with visual inspection
of positional base preference in codons, especially where there is
an increase in the observed/expected third position G + C.

CAUTION: In some cases we may not have predicted the correct
initiation codon. Where possible we choose an initiation codon
(atg, gtg, or ttg) which is preceded by an upstream ribosome
binding site sequence (optimally 5-13bp before the initiation
codon). If this cannot be identified we choose the most upstream
initiation codon.

Location/Qualifiers
1..19770

/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
/clone="I376"
57..61

/note="aggga, possible rbs upstream of Rv0654"
68..1573
/gene="Rv0654"

68..1573
/gene="Rv0654"
/note="Rv0654", (MTC1376.22), len: 501. unknown. FASTA
score: 053353 LIGNOSTILBENE-ALPHA, BETA-DIOXYGENASE (485
aa) opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity
in 523 aa overlap). Also similar to M. tuberculosis protein
MTC121C12.07c (29.5% identity in 522 aa overlap)"

/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0654"

/protein_id="CA809380.1"
/db_xref="GI:2143307"

/db_xref="SPTREMBL:O06785"

/translation="MTTAQAAESONPYLEGFLAPVSTEVATDLPVTGRIPHELDGRY
LRGNPNVAEYDPATYHWFAGDAMVHGVALRDGKARKYRNWRVTPAVCAALGEPISA
RHPDRTGIEGGPNVLTTHAGPTLALVEAGVYNYELTDELDTVGPCDFDGTUHGHT
AHPODPTGELHAYSPFARGHRVQSVIGTDGHARTVDIEVAGSPMMHSFSLTDN
YVYIDLPTFDVPMQVVPASVPRMLQPARLVITQSVLGRVRIIPDPTAALGNRMQGHSD
RLPVAWPSYPARGVNMPREGGNEDEWEDIEPCYVYHPLNAYSECRNGAEVLVDVW
RYSMFDRDGRSDRPSLDRTWLTINLATGAVTAECRDDRAQEFPRINETLVGGPHER
FAYTGVTEGFLVGAGAALSTPLYKQDCVTGSSTVASLDPLDILGEMVFPNPSARAE
DDGLMGYGNHRRGEGQLLLDAQTLESATVHLQRPVPMGFHGNWAPT"
1585. 2664
/gene="Rv0655"
CDS
1585. 2664
/note="Rv0655"
FASTA score: YRBF_ECOLI P45393 hypothetical abc
transporter atp-binding (269 aa) opt: 644 z-score: 721.8
E(): 3.4e-33 (38.5% identity in 244 aa overlap); contains
PS00017 ATP/GTP-binding site motif A, PS00211 ABC
transporters family signature, highly similar to M. leprae
MKL_MYCLE P30769 possible ribonucleotide transport atp-
(347 aa) opt: 2021 z-score:2244.4 E(): 0, (92.2% identity
in 335 aa overlap). Also similar to many other M.
tuberculosis ABC transporters eg. MTCY253.24 (33.6%
identity in 241 aa overlap)
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0655"
/protein_id="CAB09379.1"
/db_xref="GI:2143306"
/db_xref="SWISS-PROT:O06784"
/translation="MRYSDYHTGRMQPRASTEGPMPGVSVIEVNGLTKSPGSSRIWE
DVTLPITAGVSVLLGSGTGKSVFLKSLIGLLRPERGSIIDGTDIIECSAKELYEI
TLFGVLFDGALFGSMNLVDNTAFPLREHTKKKESEIRDIVNEKIALVGLGDEKFE
PGISGMRKRAGRALVLDPOIILCDPDSGLDVPRTAYLSOLIMDINAODATIL
IVTHNTNIARTVPPNNMLFRKHLVWFGPREVLLTSDPEVVRQFLNGRIGMSE
KDRATAEALLDACHAGGVEIEGVPQISATPGMPERRAVARQARVREMLHTL
PKAQAAILDLGCTHKYAVHEIGQ"
1762. 1785
/gene="Rv0655"
misc_feature
/note="PS00017 ATP/GTP-binding site motif A"
2074. 2118
/gene="Rv0655"
misc_feature
/note="PS00211 ABC transporters family signature"
complement(3052. 3435)
/gene="Rv0656c"
CDS
complement(3052. 3435)
/gene="Rv0656c"
/note="Rv0656c, (MTCI376.20, unknown), len: 127 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0656c"
/protein_id="CAB09378.1"
/db_xref="GI:2143305"
/db_xref="SPTREMBL:O06783"
/translation="MAAAVTTCTHRCLELRAAQRVAGSCEPORAEEFCRSARNADFQD
MSRMFGDVDPDVPKSVWRWIDSQHLRLARAGAVGALSVVDLLICDFAAAGLVLLH
DADYELAEHLPLDIRVRVSAADD"
complement(3530. 3685)
/gene="Rv0657c"
CDS
complement(3530. 3685)
/gene="Rv0657c"
/note="Rv0657c, (MTCI376.19), unknown, len: 51 aa; similar
to several other M. tuberculosis hypothetical proteins eg.
YW08_MYCRU Q10848 hypothetical 8.9 kd protein cy39.08c (80
aa), fasta scores; opt: 107 z-score: 182.3 E(): 0.0038,
45.8% identity in 48 aa overlap. Also similar to MTCY48_5
andAL020958|SC4H8.7 streptomycetes coellicolor cosmid 4H8 (66
aa), 41.0% identity in 39 aa overlap."
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0657c"
/protein_id="CAB09377.1"
/db_xref="GI:2143304"

/db_xref="SPTREMBL:O06782"
/translation="MSVTQIDLDDEALADVMRIAATVHTTKEAVNLAMRDYVERFERRIE
ALARSRE"
complement(3691. 3696)
/note="possible RBS upstream of Rv0657c"
complement(3761. 4477)
/gene="Rv0658c"
CDS
complement(3761. 4477)
/note="Rv0658c"
/note="Rv0658c, (MTCI376.18), len: 238, unknown,
probablemembrane protein, similar to YPRB_ECOLI P33774
hypothetical24.3 kd protein (urf 1) (217 aa), fasta
scores; opt: 174 z-score: 215.8 E(): 5.3e-05, (25.6%
identity in 223 aa overlap). Also similar to MTCY359.10
(28.7% identity in 178 aa overlap)
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0658c"
/protein_id="CAB09376.1"
/db_xref="GI:2143303"
/db_xref="SPTREMBL:O06781"
/translation="MEAGRADTAPSHRWGLGAFVLVFLVASTSLAVVLTGHPV
SAGVIALAALAPTVAAGLAILITRLRGNLRTDLRLRWSMRGLRGLMFCGCMVLT
IPASLVYTAIVGPEANSVVRIFGVRASHPWALVVLVYVFPVAPLCEEIIYRGLLMG
AVDREWGSRWALVYTVYVFAHAHEFAFAPLVLVVAIPALARTISGGLASIVTHQV
TNLLPGIVLLGLTGAIISLP"
complement(4480. 4483)
/note="possible RBS upstream of Rv0658c"
complement(4753. 5061)
/gene="Rv0659c"
CDS
complement(4753. 5061)
/gene="Rv0659c"
/note="Rv0659c, (MTCI376.17), len: 102; unknown, similar
to YW28_MYCTU Q10867 hypothetical 12.3 kd protein cy39.28
(114 aa), fasta scores; opt: 144 z-score: 213.2 E():
7.3e-05, 30.8% identity in 107 aa overlap. Also similar to
MTCY09F9.22 (32.7% identity in 101 aa overlap)
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0659c"
/protein_id="CAB09387.1"
/db_xref="GI:2143302"
/db_xref="SPTREMBL:O06780"
/translation="MRGELWFAATPGDPRVLVLTTRDPVADRIGAVVVVALTRRGG
LVSELELTAVENRVPSDCVNFNDNIHTLPRTAFRRRITRLSPARLHEACQTLRASTGC
"
complement(5048. 5293)
/gene="Rv0660c"
CDS
complement(5048. 5293)
/gene="Rv0660c"
/note="Rv0660c, (MTCI376.16), len: 81, some similarity to
IAP016485_130 Haibacterium sp: NRC-1 plasm (100 aa),
32.4%identity in 74 aa overlap"
/codon_start=1
Query Match 80.0%; Score 166.4; DB 1; Length 19770;
Best Local Similarity 87.5%; Pred. No. 2.1e-18;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAGCGCTACACCTGGCCCGCGTGGCCGCTACAAGGTCAACAAGAGCTGG 60
Db 10758 TCAAGGAGAGCGCTACACCTGGCCCGCGTGGCTATAGGTCAACAAGAGCTGG 10817
Qy 61 GCCTGAACCGCGCCAGCGGTACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 10818 GCCTGACATGTCGGGAGGCCATCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 10877
Qy 121 CCATCGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 10878 CCATCGAATATCTGCTCGCTTGCACGAGGTGCAGACCATGACCGTCCCGCGCGCG 10937
Qy 181 TCGAGGTCCTCGGTGAGGTGGACGACAT 208
||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 10938 TCGAGGTCCCGGTGAAACCGACGACAT 10965

Search completed: November 13, 2002, 01:27:11
Job time : 660.723 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic acid - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 ; Search time 1109.13 Seconds
(without alignments)
3037.202 Million cell updates/sec

Title: US-09-697-123b-17
Perfect score: 208
Sequence: 1 tcaagagagaagcgtacgac.....ccggtggaacccgacgacat 208

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estnu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_hic:*
 - 9: gb_estl:*
 - 10: gb_est2:*
 - 11: gb_hic:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estfun:*
 - 16: em_estom:*
 - 17: gb_gss:*
 - 18: em_gss_hum:*
 - 19: em_gss_inv:*
 - 20: em_gss_pln:*
 - 21: em_gss_vrt:*
 - 22: em_gss_fun:*
 - 23: em_gss_nam:*
 - 24: em_gss_mus:*
 - 25: em_gss_other:*
 - 26: em_gss_pro:*
 - 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	155.2	74.6	1282	9	AI770311
2	48.2	23.2	649	12	BG908023
3	48.2	23.2	1013	10	BE418320
C 4	48	23.1	487	12	BF444643
5	48	23.1	704	13	BM190539
6	46.6	22.4	363	12	BG907534

7	45	21.6	517	13	BJ233830
8	45	21.6	684	13	BJ258518
9	43.8	21.1	448	12	BG349727
10	43.8	21.1	468	9	AL830817
11	43.8	21.1	514	14	BQ577676
12	42.2	20.3	648	14	BQ577676
13	42.2	20.3	413	9	AL830932
14	42.2	20.3	645	10	AM056150
C 15	42	20.2	896	17	CNS03BQW
16	41.8	20.1	573	14	BQ246115
17	41.8	20.1	630	12	BG908999
18	41.8	20.1	649	14	BQ246063
C 19	41.6	20.0	512	13	BM285614
20	41.6	20.0	514	10	BE476577
C 21	41.6	20.0	534	12	BE756092
C 22	41.6	20.0	966	17	CNS02MV2
23	41.4	19.9	458	14	BQ754171
24	41.4	19.9	496	12	BG369919
25	41.4	19.9	506	10	AV835256
26	41.4	19.9	506	13	BJ469527
27	41.4	19.9	592	13	BI778452
28	41.4	19.9	799	12	BF618921
29	41.4	19.9	870	12	BF624572
30	41.2	19.8	1489	11	AI707010
31	40.8	19.6	462	9	AI676691
32	40.6	19.5	471	9	AL831237
33	40.6	19.5	485	12	BG354527
34	40.6	19.5	577	10	AW564194
35	40.4	19.4	616	13	BI722737
36	40.4	19.4	656	13	BI960536
37	40.2	19.3	512	13	BM135567
38	40.2	19.3	717	13	BJ267444
39	40	19.2	460	9	AU070110
C 40	39.8	19.1	672	12	BG416431
41	39.8	19.1	704	13	BI489596
C 42	39.8	19.1	935	10	BE196164
43	39.6	19.0	408	13	BJ470020
44	39.6	19.0	416	10	AM484285
45	39.6	19.0	420	9	AU090131

ALIGNMENTS

RESULT 1
AI770311/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

AI770311 1282 bp mRNA linear EST 24-JAN-2000
42 Mycobacterium anaerobic stationary phase library Mycobacterium smegmatis CDNA, mRNA sequence.
AI770311 GI:6742680
EST.
Mycobacterium smegmatis.
Mycobacterium smegmatis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium.
1 (bases 1 to 1282)
Murugasu-Oei, B., Tay, A. and Dick, T.
Upregulation of stress response genes and ABC transporters in anaerobic stationary-phase Mycobacterium smegmatis
Mol. Gen. Genet. 262 (4-5), 677-682 (1999)
20092472
Contact: Murugasu-Oei, B.
Mycobacterium Laboratory
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore 117609, Republic of Singapore
Tel: 65 874 3011
Fax: 65 779 1117
Email: mcbom@mcmb.nus.edu.sg
Insert Length: 1282 Std Error: 0.00
Seq primer: T3 Forward; T7 Backward.

FEATURES
source

Location/Qualifiers
1..1282

```
/organism="Mycobacterium smegmatis"
/strain="mc2155"
/db_xref="taxon:1772"
/clone_lib="Mycobacterium anaerobic stationary phase
library"
/lab_host="E. coli XL1-Blue MRF'"
/notes="Vector: Lambda ZAP II; Bacilli were disrupted using
a Mini bead beater (Biospec). RNA was isolated using the
RNeasy protocol (Qiagen). Purified RNA was subjected to 2
rounds of digestion with RNase-free DNase I (Promega).
DNase I was heat-inactivated at 75oC for 5 min. and
removed by using RNeasy columns followed by phenol
extraction and ethanol precipitation. The RNA
preparations were confirmed to be free of genomic DNA
contamination by carrying out PCR and RT-PCR using the
Access kit (Promega) and primers specific for the
histone-like protein gene hip (Lee et al., 1998). cDNA
was synthesized using random hexamer primers (Promega)
and Stratagene's cDNA synthesis kit. cDNA fragments were
ligated into lambda ZAP II vector and packaged in vitro
using Stratagene reagents."
BASE COUNT      230 a 395 c 405 g 252 t
ORIGIN

Query Match      74.6%; Score 155.2; DB 9; Length 1282;
Best Local Similarity 84.1%; Pred. No. 3.4e-25;
Matches 175; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 TCAAGGAGACCGGTACGACCTCGCGGGTGGCCGCTACAAGGTCAACAAGAGCTCG 60
Db 236 TCAAGGAGAGCGGTACGACCTCGCGGGTGGCCGCTTACAAGGTCAACAAGAGCTGG 177
Qy 61 GCCTGAACACCGCTCCCGGATACGACGACCACTCTGACCGAGAGGAGCGTGTGCGCA 120
Db 176 GCCTGAACCGGGCAACCGCATCACAGCTCGAGCTGACCGAGAGAGGAGCGTGTGCGCA 117
Qy 121 CCATCGAGTACTGGTCCGCTGCACGAGGGCCACACGATGACCGTCCCGGGCGGAG 180
Db 116 CCATCGAGTACTGGTCCGCTGCACGAGGTCAGACTTCGATGACCGTCCCGGGTGGCG 57
Qy 181 TCGAGGTGCGGGTGGAAACGCGACAT 208
Db 56 TCGAGGTTCGCGTGGAGGTGCAGACAT 29

RESULT 2
BG908023      649 bp mRNA linear EST 05-JUN-2001
LOCUS
DEFINITION
TaLr1164F05R TaLr1 Triticum aestivum cDNA clone TaLr1164F05 5',
mRNA sequence.
ACCESSION
BG908023
VERSION
BG908023.1 GI:14315686
KEYWORDS
EST.
SOURCE
bread wheat.
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 649)
Cloutier,S., Dong,G. and Walsh,A.
Wheat functional genomics - Thatcher Lr1 cDNA library
Unpublished (2001)
Contact: Dr. Sylvie Cloutier
Cereal Research Centre, Agriculture and Agri-food Canada
195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
Tel: (204) 983-2340
Fax: (204) 983-4604
Email: scloutier@em.agr.ca
was cloned directionally, not all sequences generated with reverse
primer were from the 5' end (same with forward primer and 3' end).
Average inset size is >2.2 kb
Plate: 164 row: F column: 05
Seq primer: M13 Reverse.
```

```
Location/Qualifiers
1.. 649
/organism="Triticum aestivum"
/cultivar="Thatcher Lr1"
/db_xref="taxon:4565"
/clone="TaLr1164F05"
/clone_lib="TaLr1"
/tissue_type="Leaf tissue"
/dev_stage="14 Days old"
/lab_host="E. coli XL0LR"
/notes="Vector: Lambda ZapII; mass excised in plasmid
vector pBK-CMV (Stratagene).; Site.1: ECORI; Site.2: XhoI;
mRNA obtained from wheat NIL Thatcher Lr1 24 hours after
inoculation with leaf rust pathogen puccinia tritici-na
race BBB carrying the avirulence gene Avr1."
BASE COUNT      162 a 178 c 179 g 130 t
ORIGIN

Query Match      23.2%; Score 48.2; DB 12; Length 649;
Best Local Similarity 57.7%; Pred. No. 0.28;
Matches 86; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 25 CGCGGTGGCGCTACAGGTCAACAAGAGCTCGGCTGAACACCGGTCCCGCATCA 84
Db 190 CGCGGTGGCGCTACAGGTTCGGCAAGACCGTGGCGAGGCGAGCTTCGCCAAGGTCA 249
Qy 85 CGACGACCACTCTGACCGGAAGAGGAGCGTGTGCGCCACCATCGAGTACCTGGTCCGCTGC 144
Db 250 AGATCGCCAAGGACACCCGCCACGCGGCCACCTGCGCCATCAAGGTGCTCGACCCCAACC 309
Qy 145 ACGAGGGCCACACCGAGTACCGTCCCG 173
Db 310 ACGTCTCGCGCAAGATGGTGTGAGCAG 338

RESULT 3
BG418320
LOCUS
DEFINITION
SCL024.F08R990724 ITEC SCL Wheat Leaf Library Triticum aestivum
cDNA clone SCL024.F08, mRNA sequence.
ACCESSION
BG418320
VERSION
BG418320.1 GI:9416166
KEYWORDS
EST.
SOURCE
bread wheat.
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 1013)
Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
Sorrells,M., Warburton,M. and Wenzel,G.
International Triticeae Est Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticeae
Unpublished (2000)
Contact: Cloutier S
Cereal Research Centre, Agriculture & Agri-Foods Canada
Winnipeg MT CANADA
Tel: 204 983 2340
Fax: 204 983 4604
Email: scloutier@em.agr.ca
International Triticeae Est Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.
Location/Qualifiers
1.. 1013
/organism="Triticum aestivum"
/cultivar="Thatcher Lr1"
/db_xref="taxon:4565"
/clone="SCL024.F08"
/clone_lib="ITEC SCL Wheat Leaf Library"
```

```

/tissue_type="young leaf"
/dev_stage="after 24 hour challenge with LR-AVR1"
/note="Vector: Lambda Zap; 1.0 Kbp average insert size."
BASE COUNT      246 a      271 g      223 t      36 others
ORIGIN

Query Match      23.2%; Score 48.2; DB 10; Length 1013;
Best Local Similarity 57.7%; Pred. No. 0.28;
Matches 86; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 25 CCGGGTGGCGGTACAAAGTCAACAAGAGCTCGGCTGAACACCGCGTCCCGCATCA 84
|||||
Db 186 CCGGGTGGCGGTACAAAGTCAACAAGAGCTCGGCGGAGCGCTTCCCGAAGTCA 245
|||||
QY 85 CGAGCACCCTCTGACCAAGAGAGCTCGTGGCCACATCGAGTATGCTGCTCGCTGC 144
|||||
Db 246 AGATGCCCAAGGACACCGCAACCGCCACCTCGCCATCAAGGTGCTGCGACGCAAC 305
|||||
QY 145 ACGAGGCGCACACCGATGACCTGCCG 173
|||||
Db 306 ACGTCTCCGCCACAAGATGTCGAGCAG 334
|||||

RESULT 4
BF444643/c      BF444643      487 bp      mRNA      linear      EST 01-DEC-2000
LOCUS
DEFINITION      26217 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION      BF444643
VERSION      BF444643.1 GI:11504735
KEYWORDS      EST.
SOURCE      pig.
ORGANISM      Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE      1 (bases 1 to 487)
AUTHORS      Fahrrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
TITLE      Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL      EST published (2000)
COMMENT      Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR primers
FORWARD: AGCAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCACGACG
Plate: 98 row: L column: 4
Seq primer: ATTTAGTGCACACTATAG.
FEATURES
source
1..487
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT      72 a      141 c      162 g      112 t
ORIGIN

Query Match      23.1%; Score 48; DB 12; Length 487;
Best Local Similarity 51.9%; Pred. No. 0.3;
Matches 108; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGGCTTACAAAGTCAACAGAGCTCG 60
|||||

/tissue_type="young leaf"
/dev_stage="after 24 hour challenge with LR-AVR1"
/note="Vector: Lambda Zap; 1.0 Kbp average insert size."
BASE COUNT      246 a      271 g      223 t      36 others
ORIGIN

Query Match      23.2%; Score 48.2; DB 10; Length 1013;
Best Local Similarity 57.7%; Pred. No. 0.28;
Matches 86; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 25 CCGGGTGGCGGTACAAAGTCAACAAGAGCTCGGCTGAACACCGCGTCCCGCATCA 84
|||||
Db 186 CCGGGTGGCGGTACAAAGTCAACAAGAGCTCGGCGGAGCGCTTCCCGAAGTCA 245
|||||
QY 85 CGAGCACCCTCTGACCAAGAGAGCTCGTGGCCACATCGAGTATGCTGCTCGCTGC 144
|||||
Db 246 AGATGCCCAAGGACACCGCAACCGCCACCTCGCCATCAAGGTGCTGCGACGCAAC 305
|||||
QY 145 ACGAGGCGCACACCGATGACCTGCCG 173
|||||
Db 306 ACGTCTCCGCCACAAGATGTCGAGCAG 334
|||||

RESULT 5
BM190539
LOCUS
DEFINITION      POSM0100013_D07F porcine skeletal muscle cDNA library (POSM) Sus
scrofa cDNA 5', mRNA sequence.
ACCESSION      BM190539
VERSION      BM190539.1 GI:17526502
KEYWORDS      EST.
SOURCE      pig.
ORGANISM      Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE      1 (bases 1 to 704)
AUTHORS      Yao,J.,, Coussens,P., Ernst,C. and Saama,P.
TITLE      Analysis of expressed sequence tags from a normalized porcine
skeletal muscle cDNA library
JOURNAL      Unpublished (2002)
COMMENT      Contact: Jianbo Yao
Department of Animal Science
Michigan State University
B215 Anthony Hall, East Lansing, MI, USA
Tel: 517-355-8443
Fax: 517-353-1699
Email: yaoj@msu.edu
Seq primer: M13 reverse.
FEATURES
source
1..704
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="porcine skeletal muscle cDNA library (POSM)"
/sex="male and female"
/tissue_type="skeletal muscles"
/dev_stage="45 d and 90 d of gestation, birth, 7 wk and 1
yr of age"
/lab_host="DH10B"
/note="Organ: hind limbs; Vector: pSPORT1; Site_1: NotI;
Site_2: SalI"
BASE COUNT      138 a      253 c      207 g      106 t
ORIGIN

Query Match      23.1%; Score 48; DB 13; Length 704;
Best Local Similarity 51.9%; Pred. No. 0.31;
Matches 108; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGGCTTACAAAGTCAACAGAGCTCG 60
|||||
Db 14 TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGGCTTACAAAGTCAACAGAGCTCG 73
|||||
QY 61 GCCTGAACACCGCTCCCGATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120
|||||
Db 74 TGCAGTACACCGGAGCGCAGCTGCTGCCCGCCACTCAGAACACCGGATCGCCCTGTCA 133
|||||
QY 121 CCATCTGAGTACCTCGCTCGCGCTGACCGAGGCGCCACACGATGACCGTCCCGGGCGGAG 180
|||||
Db 134 TCAGCGAGCGCGCTCAGACACCGAGGAGACACCGCCGCTGAGCGTCTGTCGCGCC 193
|||||
QY 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208
|||||

```

[illegible]

```

REFERENCE
AUTHORS      Ogihara,Y. and Murai,K.
TITLE        Expressed genes in Triticum aestivum
JOURNAL      Unpublished (2002)
COMMENT      Contact: Tadasu Shin-i
              Center For Genetic Resource Information
              National Institute of Genetics
              1111 Yata, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-559-81-6856
              Fax: 81-559-81-6855
              Email: tshini@genes.nig.ac.jp.
              Location/Qualifiers
                1. .684
                  /organism="Triticum aestivum"
                  /cultivar="Chinese Spring"
                  /db_xref="taxon:4565"
                  /clone="Whhd07"
                  /clone_lib="Y. Ogihara unpublished cDNA library, Wh_h"
                  /tissue_type="spike at heading date"
                  /dev_stage="Feekes' scale 10.5"
                  /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
                  Site1: EcoRI; Site2: XhoI; Plants were grown under
                  hydroponic conditions at UC Davis, salt stressed for 12
                  hours, and for 7 days, then dissected and frozen (Akhunov
                  in J Dvorak Lab). Total RNA was prepared from sheath
                  tissue, equal quantities of RNA were pooled from the two
                  samples, polyA was purified from the pooled RNA, a cDNA
                  library was made, and the cDNA clones were in vivo
                  excised to give pBluescript phagemids in the TJ Close lab
                  at the University of California, Riverside (Akhunov, Chin
                  , Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).
                  Plasmid DNA preparations and DNA sequencing were
                  performed in the OD Anderson lab (all other authors)."
BASE COUNT  180 a 178 c 190 g 136 t
ORIGIN
Query Match      21.6%; Score 45; DB 13; Length 684;
Best Local Similarity 56.4%; Pred. No. 1.4;
Matches 84; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY  25  CGCGGGTGGCGCTCACAGGCTCACAGAGCTCGGCTGAACACCGCGTCCCGATCA 84
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  180  CGCGGGTGGCGCTCACAGGCTCGGCAAGACCGTCCGGCGAGGCGAGCTTCGCCAAGGTCA 239
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  85  CGACGACCACTCTGACCGAAGAGGAGCTGTCGCCACCATCGAGTACCTGCTCGCGCTGC 144
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  240  AGATCCCAAGGACACCCCGACCGCGCACCTGCGCCATCAAGTGTCTGACCGCAACC 299
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  145  ACAGAGGGCCACACCGATGACCGTCCCG 173
      ||| || || || || || || || || || || || || || || || || || ||
Db  300  ACGTCTCCGGCACAAGATGTCGAGCAG 328

RESULT 9
BG349727
LOCUS      BG349727 448 bp mRNA linear EST 01-MAR-2001
DEFINITION 947030C07.y2 947 - 2 week shoot from Barkan lab Zea mays cDNA, mRNA
sequence.
ACCESSION  BG349727
VERSION     BG349727.1 GI:13178469
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 448)
Walbot,V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
JOURNAL
COMMENT    Contact: Walbot V
            Department of Biological Sciences

Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 947030 row: C column: 07.
Location/Qualifiers
  1. .448
    /organism="Zea mays"
    /cultivar="B73"
    /db_xref="taxon:4577"
    /clone_lib="947 - 2 week shoot from Barkan lab"
    /tissue_type="leaf and stem, including leaf base"
    /dev_stage="2 week old seedling (3 leaves)"
    /lab_host="XLI-Blue"
    /note="Organ: shoot; Vector: Lambda ZAP (pBluescript SK-);
    Site1: EcoRI; Site2: XhoI; Directionally cloned using
    Stratagene's Unizap XR cDNA cloning kit with the 5' end
    at the EcoRI site. The library represents 8 x 10e5
    independent recombinant phage. The plants were greenhouse
    grown."
BASE COUNT  101 a 140 c 136 g 71 t
ORIGIN
Query Match      21.1%; Score 43.8; DB 12; Length 448;
Best Local Similarity 53.9%; Pred. No. 2.6;
Matches 90; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY  25  CGCGGGTGGCGCTCACAGGCTCACAGAGCTCGGCTGAACACCGCGTCCCGATCA 84
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  168  CGCGGGTGGCGCTCACAGGCTCGGCAAGACCATCGGCGAGGCGAGCTTCGCCAAGGTCA 227
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY  85  CGACGACCACTCTGACCGAAGAGGAGCTGTCGCCACCATCGAGTACCTGCTCGCGCTGC 144
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  228  AGCACGCGCGGACTCCCGCACCGCGCGCTCCGCCCATCAAGTGTCTGACCGCAACC 287
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  145  ACAGAGGGCCACACCGATGACCGTCCCGGGCGGAGTCGAGGTCCCG 191
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  288  ACGTCTCCGCCACAGGATGTCGAGCAGATCAAGCGGAGATTTTCG 334

RESULT 10
AL830817
LOCUS      AL830817 468 bp mRNA linear EST 16-JUL-2002
DEFINITION  AL830817 q:242 Triticum aestivum cDNA clone E02_q242_plate_12, mRNA
sequence.
ACCESSION  AL830817
VERSION     AL830817.1 GI:21842597
KEYWORDS   EST.
SOURCE     bread wheat.
ORGANISM   Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
clade; Triticeae; Triticum.
1 (bases 1 to 468)
Wilson,I., Beswick,R., Shepherd,S., Barker,G., Parker,J., Owen,P.,
Edwards,D., Coghill,J., Holdsworth,M., Lenton,J., Shewry,P. and
Edwards,K.
A BBSRC-funded wheat EST resource for the academic community
Unpublished (2002)
Contact: Barker G
Institute of Arable Crop Research
Long Ashton, Bristol BS41 9AF United Kingdom.
Location/Qualifiers
  1. .468
    /organism="Triticum aestivum"
    /cultivar="florida"
    /db_xref="taxon:4565"
    /clone="E02_q242_plate_12"
    /clone_lib="q:242"
    /tissue_type="egg cell"
    /dev_stage="pre-fertilisation"
BASE COUNT  92 a 150 c 139 g 87 t

```


ORIGIN

Query Hatch	21.13;	Score 43.8;	DB 9;	Length 468;
Best Local Similarity	50.7%;	Pred. No. 2.6;		
Matches 105;	Conservative 0;	Mismatches 102;	Indels 0;	Gaps 0;
QY 2	CAAGGAGAAGCCTACGACCCTCGCGGGTGGCCGCTACAAGGTCAACAGAAAGTCGG	61		
Db				
140	CGACGAGACCGGGTAGCTAAGGGGCTCGTCACTATAGTACGGACGGGTGA	199		
QY 62	CTGTGAACACCGGTCGCCGATTACGAGCACCATCTTGACCGAAGAGAGCTGTGCGCCAC	121		
Db				
200	GGTGACCCCCATGTGCGGCATCTCCAGCATCACCCCTCATCAACGAGTTACAGGTCTACAA	259		
QY 122	CATCGAGTACCTGGTCCGCTTGCACGAGGGCCACACCAGCATCACCGTCCCGGGCGGAGT	181		
Db				
260	GGACGTGGAGCTGCCGAGAGAAGTTCTCAGCGGTACGGTGGACCGAGGGCTCGGGCTCTCT	319		
QY 182	CGAGGTGCCGGTGGAAACCGACAGCAT	208		
Db				
320	CAAGTGGCACTGGGCTCCGACACCGT	346		

RESULT 11	BQ577676	514 bp	linear	EST 19-JUN-2002		
LOCUS	3524_l_39_l_H12_y_l_3524	Mature pollen from Sheila McCormick's				
DEFINITION	lab Zea mays cDNA, mRNA sequence.					
ACCESSION	BQ577676					
VERSION	BQ577676.1	GI:21480993				
KEYWORDS	EST.					
SOURCE	Zea mays.					
ORGANISM	Zea mays					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.					
REFERENCE	1 (bases 1 to 514)					
AUTHORS	Walbot,V.					
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford University					
JOURNAL	Unpublished (1999)					
COMMENT	Contact: Walbot V					
	Department of Biological Sciences					
	Stanford University					
	855 California Ave, Palo Alto, CA 94304, USA					
	Tel: 650 723 2327					
	Fax: 650 725 8221					
	Email: walbot@stanford.edu					
	Plate: 3524_l_39_l row: H	column: 12.				

```

FEATURES
  source
    Location/Qualifiers
      1..514
        /organism="zea mays"
        /cultivar="B73"
        /db_xref="taxon:4577"
        /clone_lib="3524 - Mature pollen from Sheila McCormick's lab"
        /tissue_type="pollen"
        /dev_stage="mature"
        /lab_host="SOLR"
        /note="Vector: Stratagene's Uni-Zap XR (pBluescript SK-); Site1: EcoRI (5-prime); Site2: XhoI (3-prime); Unamplified cDNA library directionally cloned by Rima Kuitkauskas using Stratagene's Uni-Zap system. Insert sizes ranged from 0.5Kb to 2Kb. 50 microliter aliquot had 338,000 pfu when it was made in Sept, 1995, from Oligo dT-primed poly A+ RNA."
      125 a 146 c 153 q 90 t

```

BASE COUNT	125 a	146 c	153 g	90 t
ORIGIN				
Query Match	21.1%; Score 43.8; DB 14; Length 514;			
Best Local Similarity	53.9%; Pred. No. 2.7;			
Matches	90; Conservative	0; Mismatches	77; Indels	0; Gaps 0;

QY	25	CGCGGGTGGGCGGTACAAAGCTCAACAGAAGTCTGGCCCTGAACACCGCGTCCCGCATCA	84
Db	161	CGCGGGTGGGCGGTACGAGTTCGGCAAGACATCGCGGAGGCGAGCTTCGCCAAGGTCA	220
QY	85	CGACGACCACCTCTGACCGAAGAGGAGCTGCTGCGCCACCATCGAGTACCTGGTCCGCGTGC	144
Db	221	AGCAGCGCGCGACTCCGCGACCGCGCGCTCCGCGCCATCAAGGTGCTCGACCGCAACC	280
QY	145	ACGAGGGGCCACACGATGACCGTCCCGGCGGAGTCGAGGTGCCG	191
Db	281	ACGTCTCCGCCACAGGATGTCGAGCAGATCAAGCGGAGATTTCG	327
RESULT	12		
LOCUS	BQ579756		
DEFINITION	WHE2974_B09_D182s Wheat dormant embryo cDNA library	648 bp	mRNA linear EST 19-JUN-2002
ACCESSION	BQ579756		
VERSION	BQ579756.1		
KEYWORDS	EST.		
SOURCE	Bread wheat.		
ORGANISM	Triticum aestivum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae		
AUTHORS	1 (bases 1 to 648) Anderson,O.D., Chao,S., Chin,A., Close,T.J., Doherty,L., Fenton,R.D., Lazo,G.R., Rausch,C.J., Walker-Simmons,M.K. and Wilson,C.		
TITLE	The structure and function of the expressed portion of the wheat genomes - Dormant embryo cDNA library		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Olin Anderson US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center 800 Buchanan Street, Albany, CA 94710, USA Tel: 5105595773 Fax: 5105595818 Email: oandern@pw.usda.gov Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20 Seq primer: SK primer.		

```

FEATURES
source
seq primer: SK primer.
Location/Qualifiers
1..648
/organism="Triticum aestivum"
/cultivar="Brevor"
/db_xref="taxon:4565"
/clone="WHE2974_B09_D18"
/clone_lib="Wheat dormant
/tissue_type="Seed embryo"
/dev_stage="Mature seed"
/lab_host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown to seed
maturity under conditions favoring seed dormancy (L.
Doherty at K. Walker-Simmons lab, Washington State
University, Pullman, WA). Embryos were cut from mature
dormant seed (Doherty). Total RNA was prepared from these
embryos, polyA was purified, a cDNA library was made, and
the cDNA clones were in vivo excised to give pBluescript
phagemids in the TJ clone lab at the University of
California, Riverside (Chin, Fenton). Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
148 a 185 c 221 g 94 t

```

```

Query Match      20.7%; Score 43; DB 14; Length 648;
Best Local Similarity 52.5%; Pred. No. 4;
Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY      28 GGGTGGGCGGTACAAAGGTCACAGAGCTGGCGCTGAACACCGCGTCCCGATCAGCA 87
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 196 GCGGCGCGCGGGAGGTCCTGCACACGAGGCGCGCGCTGCCGTACGGGCCCGGGA 255
QY 88 CGACCACTCTGACCGAAGAGGAGTGTCTGCCACCATCGAGTACCTGTCGCCCTGCACG 147
    ||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 256 CGATGGCGCTGATGGGGTTGCGGCATGTCGGGCCCATCGCTACCTGTGTGTACCAGA 315
    ||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 148 AGGCCACACACCATGACGCTCCCGGGCGGAGTCGAGGTGCGGTGGAACCGACGAC 206
    ||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 316 AGCGAGGCCCGCGCGCGCCACGAGGTGCGCAAGGTGCGCGTCGGACACGCGGAC 374

RESULT 13
AL830932 413 bp mRNA linear EST 16-JUL-2002
LOCUS AL830932 q:242 Triticum aestivum cDNA clone C04_q242_plate_15, mRNA
DEFINITION sequence.
ACCESSION AL830932
VERSION AL830932.1 GI:21842712
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
REFERENCE 1 (bases 1 to 413)
AUTHORS Wilson,I., Beswick,R., Shepherd,S., Barker,G., Parker,J., Owen,P.,
Edwards,D., Coghill,J., Holdsworth,M., Lenton,J., Shewry,P. and
Edwards,K.
TITLE A BSRG-funded wheat EST resource for the academic community
JOURNAL Unpublished (2002)
COMMENT Contact: Barker G
Institute of Arable Crop Research
Long Ashton, Bristol BS41 9AF United Kingdom.
FEATURES
Source Location/Qualifiers
1..413
/organism="Triticum aestivum"
/cultivar="florida"
/db_xref="taxon:4565"
/clone="C04_q242_plate_15"
/clone_lib="q:242"
/tissue_type="egg cell"
/dev_stage="pre-fertilisation"
BASE COUNT 83 a 129 c 123 g 77 t 1 others
ORIGIN

Query Match 20.3%; Score 42.2; DB 9; Length 413;
Best Local Similarity 51.9%; Pred. No. 6;
Matches 95; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 26 GCGGCTGGCGCGCTCAAGAGTCAACAAGAGCTCGGCTGAACACCGCGTCCCGCATCAC 85
    ||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 106 GGGGCTCGTCACTACATTGTCACGAGCGGCTGGAGGTGACCCCATGTCCGCCATCTC 165
    ||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 86 GACGACCACTGTACCGAAGAGAGGTCGTGCGCCACCATCGAGTACCTGGTCCGCTGCA 145
    ||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 166 CAGCATCACCTCTCATCAACAAGTTTCAGCGTCGACAGAGAGCTGGAGTCCCGAGAAGTT 225
    ||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 146 CGAGGCGCACACCATGACGTCGCCGGCGGAGTCGAGGTGCGGTGGAACCCACCA 205
    ||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 226 CGTCAGCTCGGCATGACGAGGGGCTCGGCTCCTCGAGGTGGCACTGGGCTCCGACAC 285
    ||| |||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 206 CAT 208
    |||
Db 286 CGT 288
    |||

RESULT 14
AW056150 645 bp mRNA linear EST 27-SEP-1999
LOCUS AW056150 66004G09.y1 660 - Mixed stages of anther and pollen Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW056150
VERSION AW056150.1 GI:5928858

```

```

KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 645)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660004 row: G column: 09.
FEATURES
Source Location/Qualifiers
1..645
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XL0LR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 121 a 200 c 221 g 103 t
ORIGIN

Query Match 20.3%; Score 42.2; DB 10; Length 645;
Best Local Similarity 57.0%; Pred. No. 6.1;
Matches 77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 58 TCGGCTGAACACGCGTCCCGCATCAGCAGCACCTCTGACCGAAGAGAGCGTGTGTCG 117
    |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 TCGGCTGCCATCTGGCGCGCGGAGATCGCGCTCGACGACACAGCGTGTGTCG 112
    |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 118 CCACCATCGATCTGTGTCGCTGCAGGAGGCCACACGATGACCGTCCCGGGCG 177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 GCACATTCGGGTACTGGCGCGCGGAGTACTTCATGTACGGAGAGGTGACCGACAGGTCG 172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 178 GAGTCGAGTCCCG 192
    ||| ||| ||| ||| |||
Db 173 ACGTACGCGTTGG 187
    ||| ||| ||| ||| |||

RESULT 15
CNS03BQW 896 bp DNA linear GSS 15-MAY-2000
LOCUS CNS03BQW/c Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
DEFINITION 012P03 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL236849
VERSION AL236849.1 GI:7895984
KEYWORDS GSS: genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 896)
AUTHORS Roest-Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence

```

Search completed: November 13, 2002, 04:00:14
Job time : 1120.63 secs

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	81.6	39.2	5096	10	US-09-984-711-5		Sequence 5, Appli
2	81.6	39.2	5099	9	US-10-075-460-5		Sequence 5, Appli
3	81.6	39.2	5099	10	US-09-887-052-1		Sequence 1, Appli
4	81.6	39.2	5099	10	US-09-887-052-3		Sequence 3, Appli
5	81.6	39.2	5099	10	US-09-887-052-5		Sequence 5, Appli
6	38	18.3	720	9	US-09-999-745-14		Sequence 14, Appli
7	37	17.8	4039	10	US-09-205-448-7		Sequence 7, Appli
8	36.8	17.7	315	10	US-09-799-514-7		Sequence 7, Appli
9	36.4	17.5	315	10	US-09-923-876-3629		Sequence 3629, Ap
10	36.4	17.5	394	10	US-09-360-352-9522		Sequence 9522, Ap
11	36.4	17.5	1665	10	US-09-815-242-7935		Sequence 7935, Ap
12	36.4	17.5	2329	10	US-09-816-828-9		Sequence 9, Appli
13	36.2	17.4	1704	10	US-09-815-242-7788		Sequence 7788, Ap
14	36.2	17.4	3300	10	US-09-379-931-6		Sequence 6, Appli
15	35.8	17.2	639	10	US-09-815-242-7939		Sequence 7939, Ap
16	35.8	17.2	2163	10	US-09-839-408-29		Sequence 29, Appli
17	35.8	17.2	13842	10	US-09-861-289-30		Sequence 30, Appli
18	35.8	17.2	12778	10	US-09-861-289-5		Sequence 5, Appli
19	35.4	17.0	3600	10	US-09-815-242-7692		Sequence 7692, Ap

Qy	129	TACCTGTCCGCTGCACGAGGCCACACCGATGACGCTCCCGGGGGGATCGAGGTG	188
Db	1695	TACCTGTGCGCTGCACGAGGTGAGCGCGTCATGACTTCTCCAATGGTGAAGATC	1754
Qy	189	CCGGTGGAAACCGACGACAT	208
Db	1755	CCAGTCGAGACCGATGACAT	1774

```

RESULT 2
US-10-075-460-5
: Sequence 5, Application US/10075460
: Patent No. US2020215557A1
: GENERAL INFORMATION:
: APPLICANT: MCKEL, BETTINA
: APPLICANT: BATHÉ, STEFAN
: APPLICANT: HANS, STEFAN
: APPLICANT: KREUTZER, CAROLINE
: APPLICANT: HERMANN, THOMAS
: APPLICANT: PFEFFERLE, WALTER
: APPLICANT: BINDER, MICHAEL
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCES
: FILE REFERENCE: 21847205X
: CURRENT APPLICATION NUMBER: US/10/075,460
: PRIOR FILING DATE: 2002-02-15
: PRIOR APPLICATION NUMBER: DE 10107230.9
: PRIOR FILING DATE: 2001-02-16
: PRIOR APPLICATION NUMBER: DE 10162386.0
: PRIOR FILING DATE: 2001-12-19
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: Patent version 3.1
: SEQ ID NO 5
: LENGTH: 5099
: TYPE: DNA
: ORGANISM: Corynebacterium glutamicum
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (702)..(4196)
: OTHER INFORMATION:
US-10-075-460-5

```

Query Match	39.2%	Score 81.6	DB 9	Length 5099
Best Local Similarity	66.5%	Pred. No. 1.6e-12		
Matches 133	Conservative 0	Mismatches 64	Indels 3	Gaps 1
QY	9	AAGCGCTACGACCTCGCGCGGTGGCGGCTACAAAGGTCAACAAGAGCTCGCGCCTGAAC	68	
Db	1578	AAGCGCTACGACCTCGCTGGTGGTTCGTTTACAAGATCAACCGAAGTTCGCGCCTTGGT	1637	
QY	69	ACCGCGTCCCGGATCAGCAGGACCACCTCTGACCGAAGAGGAGCCTCTGCGCCACCATCGAG	128	
Db	1638	GGCGACCACGATGGTTTGATG---ACTCTACTGAAGAGGACATCGCAACCAACCATCGAG	1694	
QY	129	TACCTGGTCCGCTCGACGAGGGCCACACCAAGTACCGTCCCGGGCGGAGTCGAGGTG	188	
Db	1695	TACCTGGTGGCTCTGCACGCGAGGTGCGGCGTATGACTTCTCCAATGGTGAAGAGATC	1754	
QY	189	CCGGTGGAAACCGAGCAT	208	
Db	1755	CCAGTCGAGACCGATGACAT	1774	

```

RESULT 3
US-09-887-052-1
: Sequence 1, Application US/09887052
: Patent No. US20020119537A1
: GENERAL INFORMATION:
: APPLICANT: MOECKEL, Bettina
: APPLICANT: BATHE, Brigitte
: APPLICANT: HERMANN, Thomas
: APPLICANT: PFEFFERLE, Walter
: APPLICANT: BINDER, Michael
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE

```

```

; FILE REFERENCE: 204212USOX
; CURRENT APPLICATION NUMBER: US/09/887.052
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: DE10107229.5
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
US-09-887-052-1

```

Query Match	39.2%	Score 81.6;	DB 10;	Length 5099;
Best Local Similarity	66.5%;	Pred. No. 1.6e-12;		
Matches 133;	Conservative 0;	Mismatches 64;	Indels 3;	Gaps 1;
QY	9	AAGCGCTACGACCTCGCGCGGTGGCGCGGTACAAAGTCAACAAAGTCGCGCCTGAAC	68	
Db	1578	AAGCGCTACGACCTCGCGGTGGCGGTACAAAGTCAACCGAAGTCGCGCCTGGT	1637	
QY	69	ACCGCGTCCCGGATCACGACGACACTCTGACCGAAGAGGACGCTGCTCGCCACCATCGAG	128	
Db	1638	GCGACCAACGATGTTTGATG---ACTCTTACTGAAGAGGACATCGCAACCATCGAG	1694	
QY	129	TACCTGTGTCGCGCTGCAGAGGGCCACACACGATGACCTCCCGGGCGGATCGAGGTG	188	
Db	1695	TACCTGTGCGTCTGCAGCGAGGTGCGCGGTATGACTTCTCCAATGGTGAAGAGATC	1754	
QY	189	CCGGTGGAAACCGACGACAT	208	
Db	1755	CCAGTCGAGACCGATGACAT	1774	

```

RESULT 4
US-09-887-052-3
; Sequence 3, Application US/09887052
; Patent No. US20020119537A1
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: BATHE, Brigitte
; APPLICANT: HERMANN, Thomas
; APPLICANT: PFEFFERLE, Walter
; APPLICANT: BINDER, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE TP0B GENE
; FILE REFERENCE: 204212USOX
; CURRENT APPLICATION NUMBER: US/09/887,052
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: DE10107229.5
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
US-09-887-052-3

```

	Query Match	39.2%	Score 81.6;	DB 10;	Length 5099;
	Best Local Similarity	66.5%;			
	Matches 133;	Conservative	0;	Mismatches 64;	Indels 3; Gaps 1;
QY	9	AAGCGCTACGACTTCGGCGGGTGGCCGTACAAGGTCAACAAGAAGTCGCGCCTGAAC	68		
DB	1578	AAGCGCTACGACCTTGGCTGGTGGTTCGTACAAGATCAACGCCAAGTCGCGCTTGCT	1637		
QY	69	ACCGGCTCCCGGATCACGAGCAGCCACTCTTACCGAAGAGGAGCGTCGTCGCCACCATTGAG	128		

Db 1638 GCGGACACGATGTTGATG---ACTCTTACTGAAGAGGACATCGCAACACCATCGAG 1694
QY 129 TACCTGGTCCGCTGCACGAGGGCCACACGATGACCGTCCCGGGCGGAGTCGAGGTG 188
Db 1695 TACCTGGTCCGCTGCACGAGGGCCACGATGACCTTCCCAATGGTGAAGATC 1754
QY 189 CCGGTGGAACCGAGACAT 208
Db 1755 CCAGTCGAGACCGATGACAT 1774

RESULT 5

US-09-887-052-5
; Sequence 5, Application US/09887052
; Patent No. US20020119537A1
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: BATHE, Brigitte
; APPLICANT: HERMANN, Thomas
; APPLICANT: PFEFFERLE, Walter
; APPLICANT: BINDER, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE
; FILE REFERENCE: 204212US0X
; CURRENT APPLICATION NUMBER: US/09/887,052
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: DE10107229.5
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
US-09-887-052-5

Query Match 39.2%; Score 81.6; DB 10; Length 5099;
Best Local Similarity 66.5%; Pred. No. 1.6e-12;
Matches 133; Conservative 0; Mismatches 64; Indels 3; Gaps 1;
QY 9 AAGCGCTACGACCTCGCGCGGTGGCGGTACAAAGGTCAACAAAGCTCGCCTGAAC 68
Db 1578 AAGCGCTACGACCTCGCGGTGGCGGTACAAAGGTCAACAAAGCTCGCCTGAAC 1637
QY 69 ACCGCGTCCCGGATCAGCAGCACCCTCTGACCGAAGAGACGTCGTCGCCACCATCGAG 128
Db 1638 GCGGACACGATGTTTGTATG---ACTCTTACTGAAGAGGACATCGCAACACCATCGAG 1694
QY 129 TACCTGGTCCGCTGCACGAGGGCCACACGATGACCGTCCCGGGCGGAGTCGAGGTG 188
Db 1695 TACCTGGTCCGCTGCACGAGGGCCACGATGACCTTCCCAATGGTGAAGATC 1754
QY 189 CCGGTGGAACCGAGACAT 208
Db 1755 CCAGTCGAGACCGATGACAT 1774

RESULT 6

US-09-999-745-14
; Sequence 14, Application US/09999745
; Patent No. US20020157120A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
; FILE REFERENCE: REGEN1470-1
; CURRENT APPLICATION NUMBER: US/09/999,745
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 09/316,920

; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Aequorea victoria
US-09-999-745-14

Query Match 18.3%; Score 38; DB 9; Length 720;
Best Local Similarity 49.0%; Pred. No. 0.16;
Matches 101; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGCTTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCG 60
Db 392 TCAAGGAGGAGCGCAACATCTCTGGGCAACAGCTGGAGTACACTACATCAGCCACAACG 451
QY 61 GCCTGAACACCGCTCCCGGATCAGCAGGACCACTCTGACCGAAGAGGAGCTCGTCGCCA 120
Db 452 TCTATATCACCCCGCACAAAGCAGAAACGCGCATCAAGGCCCACTTCAAGATCCGCCACA 511
QY 121 CCATCGACTACTGTGTCGGCTGCAGGAGGCGCCACACGATGACCGTCCCGGGCGGAG 180
Db 512 ACATCGAGCGGAGCGTGCAGCTCGCGGACCACTACCAGCAGAACACCCCATCGCGG 571
QY 181 TCGAGGTCCCGGTGGAACCGACGAC 206
Db 572 ACGGCCCGGTGCTGCTCCCGCAAC 597

RESULT 7

US-09-205-448-7
; Sequence 7, Application US/09205448
; Patent No. US20020137661A1
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances
; APPLICANT: Shao, Zhixin
; APPLICANT: Volkov, Alexander
; APPLICANT: California Institute of Technology
; TITLE OF INVENTION: Method for Creating Polynucleotide and Polypeptide
; FILE REFERENCE: 018097-025710PC
; CURRENT APPLICATION NUMBER: US/09/205,448
; CURRENT FILING DATE: 1998-12-04
; PRIOR APPLICATION NUMBER: US 60/067,908
; PRIOR FILING DATE: 1997-12-08
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 4039
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Actinoplanes
; OTHER INFORMATION: utahensis echinocandin B (ECB) deacylase gene
; OTHER INFORMATION: mutant M-15 created by in vitro heteroduplex
; NAME/KEY: CDS
; LOCATION: (1196)..(3559)
US-09-205-448-7

Query Match 17.8%; Score 37; DB 10; Length 4039;
Best Local Similarity 52.2%; Pred. No. 0.33;
Matches 82; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 20 CCTCGCGCGGTGGCGGCTTACAAAGGTCAACAAGGTTCGGCTGAACACCGCGTCCCG 79
Db 2050 CATCGGGATCGGCAACAACCGCAGCGTTCGGAGCCACACCGTCTCCACCGCCCGCGG 2109
QY 80 GATCAGCAGCACCACTCTGACCGAAGAGGAGCGTGTGCGCCACCATCGAGTACCTGGTCCG 139
Db 2110 GTTCGTGTGCGACCGCTGAGCCTCTGTGCGCGGCGGACCCCACTCTCTATTAGTGTGACGG 2169


```

; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 7935
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1665)
US-09-815-242-7935

Query Match      17.5%;      Score 36.4;      DB 10;      Length 1665;
Best Local Similarity 49.5%;      Pred. No. 0.43;
Matches 94;      Conservative 0;      Mismatches 96;      Indels 0;      Gaps 0;

Qy 13  GCTAGGACCTCGCGGGGTGGCGCGCTACAAGGTCAACAAGAGCTCGGCCCTGAACACCG 72
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182  GCGAGGCGCGCGCCCTGCGGGGATCAACGTCGGCTACCTCGCGAGGAGCCCAAGCTGG 241
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 73  CGTCCCGGATCACGACGACCACCTCTGACCGAAGAGGAGCTCGTCCGCCACCATCGAGTACC 132
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242  ACCCGAGGCCACCGTCGGCGACATCTGCGAAGAGCGCTCGGCAGATCAAGCAGGCC 301
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 133  TGGTTCGGCTCGCAGAGGGCCACACCGATGATGACCTCCGGGGGGAGTCCAGGTGCGCGG 192
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302  AGGCGCGGCTGGACGAGTCTATGCCGCCCTATGCCGCGGATGCCGACTTCGACGCCCC 361
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 193  TGGAAACCGA 202
      ||| ||| |||
Db 362  TGGCGGCCGA 371

RESULT 12
US-09-816-828-9
; Sequence 9, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhao, Qing A.
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 79ICP2E
; CURRENT APPLICATION NUMBER: US/09/816,828
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: pt_FL_genes Version 2.0

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:56:36 ; Search time 27.8936 Seconds
(without alignments)
2286.857 Million cell updates/sec

Title: US-09-697-123B-17
Perfect score: 208
Sequence: 1 tcaggagagcgctacgac.....ccggtggaaccgacgacat 208

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A-COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B-COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A-COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B-COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS-COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	160	76.9	970	1	US-08-250-030-1
2	160	76.9	970	5	PCT-US95-06790-1
3	144	69.2	3447	2	US-08-313-185-57
4	144	69.2	3447	3	US-09-082-614A-57
5	38.6	18.6	1575	1	US-07-988-260B-2
6	38.2	18.4	2122	4	US-09-029-603-1
7	38.2	18.4	44377	2	US-08-804-227C-7
8	38.2	18.4	44377	2	US-08-804-198-1
9	38	18.3	717	4	US-09-513-783A-49
10	38	18.3	720	3	US-09-094-359-7
11	38	18.3	720	3	US-09-172-063-13
12	38	18.3	720	4	US-09-316-919-14
13	38	18.3	768	3	US-09-094-359-11
14	38	18.3	972	3	US-09-172-063-27
15	38	18.3	1623	4	US-09-513-783A-33
16	38	18.3	2742	4	US-09-232-468A-1
17	38	18.3	4897	6	5196516-7
18	37.2	17.9	942	3	US-08-732-412-1
19	37	17.8	2929	4	US-09-543-084A-31
20	37	17.8	3193	4	US-09-543-084A-30
21	37	17.8	3222	4	US-09-543-084A-29
22	37	17.8	4050	4	US-09-543-084A-26
23	37	17.8	4093	4	US-09-543-084A-28
24	36.4	17.5	4101	4	US-09-543-084A-27
25	36.2	17.4	1153	4	US-09-372-448A-5
26	36.2	17.4	1894	4	US-09-329-350-32
27	36.2	17.4	3300	1	US-08-194-290-6

28	36.2	17.4	3300	2	US-08-614-377A-6	Sequence 6, Appli
29	36.2	17.4	3300	4	US-09-142-648B-6	Sequence 6, Appli
30	36.2	17.4	4403765	4	US-09-103-840A-2	Sequence 1, Appli
31	36.2	17.4	4411529	4	US-09-103-840A-1	Sequence 1, Appli
32	36	17.3	765	3	US-08-718-904-79	Sequence 79, Appli
33	35.8	17.2	13842	4	US-09-105-537-30	Sequence 30, Appli
34	35.8	17.2	36778	4	US-09-105-537-5	Sequence 5, Appli
35	35.8	17.2	38506	3	US-09-320-878-19	Sequence 19, Appli
36	35.4	17.0	6085	4	US-09-029-603-4	Sequence 4, Appli
37	35.2	16.9	1095	2	US-08-851-088-3	Sequence 3, Appli
38	35.2	16.9	5535	1	US-08-089-755A-1	Sequence 1, Appli
39	35.2	16.9	5535	1	US-08-089-755A-4	Sequence 1, Appli
40	35.2	16.9	5535	1	US-08-421-754-1	Sequence 1, Appli
41	35.2	16.9	5535	2	US-08-421-754-4	Sequence 4, Appli
42	35.2	16.9	5535	2	US-08-421-791-1	Sequence 1, Appli
43	35.2	16.9	5535	2	US-08-421-791-4	Sequence 1, Appli
44	35	16.8	5011	1	US-08-141-893-1	Sequence 1, Appli
45	35	16.8	5011	1	US-08-463-092B-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-250-030-1
; Sequence 1, Application US/08250030
; Patent No. 5643723
; GENERAL INFORMATION:
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: Detection of a Genetic Locus Encoding
; TITLE OF INVENTION: Resistance to Rifampin in Mycobacterial Cultures and in
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/250,030
; FILING DATE: 26-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Muetting, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 150.105US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-250-030-1

Query Match 76.9%; Score 160; DB 1; Length 970;
Best Local Similarity 85.6%; Pred. No. 5e-30;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGACTCGCGCGGTGGCGCTACAGGTCAACAAGAGCTCG 60
Db 26 TCAAGGAGAGCGCTACGACTCGCGCGGTGGCGCTACAGGTCAACAAGAGCTCG 85

Sequence 57, Application US/09082614A
Patent No. 6124098
GENERAL INFORMATION:
APPLICANT: Heym, Beate
APPLICANT: Cole, Stewart
APPLICANT: Young, Douglas
APPLICANT: Zhang, Ying
APPLICANT: Honore, Nadine
APPLICANT: Telenti, Anello
APPLICANT: Bodmer, Thomas
TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
TITLE OF INVENTION: in Mycobacterium Tuberculosis
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,614A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/313,185
FILING DATE: 12-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 02356.0068-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 3447 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-082-614A-57

Query Match 69.2%; Score 144; DB 3; Length 3447;
Best Local Similarity 80.8%; Pred. No. 3.6e-26;
Matches 168; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 1 TCAAGGAGAACCGGTACGACCTCGCGGGTGGCCCTACAAGGTCACAAAGAGCTCG 60
Db 809 TCAAGGAGAACCGGTACGACCTCGCGGGTGGCCCTACAAGGTCACAAAGAGCTCG 60
QY 61 GCCTGAACACCGCTCCCGCATACGACGACCACTCTGACCGAAGAGGAGCTCGTCCCA 120
Db 869 GTTTGCACCGCGGTGAGTTGATACGTCGTCACCGTACCGAAGAGGATGTGTCGCCA 928
QY 121 CCATCGAGTACCTGGTCCGCTGCACGAGGGCCACACGATCACCGTCCCGGGGGAG 180
Db 929 CCATAGAGTACCTGGTCCGTCGATGAGGTCAGTCGACAATGACTGTCCCGAGGTGGG 988
QY 181 TCGAGTCCCGGTGGAACGACACAT 208
Db 989 TAGAAGTGCCAGTGAAACTGACGATAT 1016

RESULT 5
US-07-988-260B-2
Sequence 2, Application US/07988260B

Patent No. 5482843
GENERAL INFORMATION:
APPLICANT: BRZEZINSKI, RYSZARD
TITLE OF INVENTION: ENZYME OF USE IN CHITOSAN
TITLE OF INVENTION: HYDROLYSIS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1100 New York Avenue, N.W., 9th Floor
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 720 kb diskette
COMPUTER: IBM PS/2, Model 30
OPERATING SYSTEM: PC-DOS 3.30
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/988,260B
FILING DATE: 19921214
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hymo, Lawrence A.
REGISTRATION NUMBER: 19,057
REFERENCE/DOCKET NUMBER: LAH/3122/98214/MJW
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1575 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE: Streptomyces N174
US-07-988-260B-2

Query Match 18.6%; Score 38.6; DB 1; Length 1575;
Best Local Similarity 51.4%; Pred. No. 0.41;
Matches 89; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 14 CTACGACCTCGCGGGTGGCGGTACAGGTCAACAGAGCTCGGCTGAACACCGC 73
Db 776 CGACGCCATCGTGATGCACGCCGCCGCCGCAACGACCGACGCTTCGGTGGCATCCGAA 835
QY 74 GTCCCGCATCAGCAGCACCCTCTGACCGAAGAGGAGCTGCTGCCACCATCGACTCT 133
Db 836 GACCGCATGAGAGAGGCCAGGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCCAA 895
QY 134 GGTCCGCTTCACGAGGGGCCACACGATGACCGTCCCGGGCGGAGTCGAGG 186
Db 896 CGCCTTCCTCGACGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGAG 948

RESULT 6
US-09-029-603-1
Sequence 1, Application US/09029603
Patent No. 6210935
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Engel, Natalie
APPLICANT: Bietenhader, Jurg
APPLICANT: Toupet, Christine
APPLICANT: Pospiech, Andreas
TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters
FILE REFERENCE: 4-20555/A/PCT
CURRENT APPLICATION NUMBER: US/09/029,603
CURRENT FILING DATE: 1998-03-20
EARLIER APPLICATION NUMBER: PCT/EP96/03643
EARLIER FILING DATE: 1996-08-19

; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; TYPE: DNA
; LENGTH: 2122
; ORGANISM: Streptomyces longisporoflavus
; FEATURE:
; NAME/KEY: misc_RNA
; LOCATION: (1)..(2122)
; OTHER INFORMATION: product = 2.1 kb region
US-09-029-603-1

Query Match 18.4%; Score 38.2; DB 4; Length 2122;
Best Local Similarity 50.8%; Pred. No. 0.52;
Matches 91; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 8 GAAGCGCTACGACCTCGCGCGGTGGCGGTACAAAGGTCAACAGAGCTCGCGCTCAA 67
Db 1470 GCACCGCTCGGGCTCGCGCTCGCGGAGATCTGACGTCAACGAGACAGACGCTGCG 1529
QY 68 CACCGCGTCCCGATCAGCAGACACACTCTGACCGAAGAGGACGTCTGCGCCACATCGA 127
Db 1530 CCTCGCCGACGAGATCGCGGCTCGCGCGGTGAGGAGCGCCCTGGCCATGGACGA 1589
QY 128 GTACCTGTGCGCTCGCAGAGGGCCACACCATGACCGTCCCGGGCGGAGTCTGAGG 186
Db 1590 GGGCAACTTGGCTTTCGGCGGACGACTCTTCAAGCGCTCCGACCTGGCGGCGTCTGACG 1648

RESULT 7
US-08-804-227C-7
; Sequence 7, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 350..14002
; FEATURE:

; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31329..36071
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36155..41830
US-08-804-227C-7

Query Match 18.4%; Score 38.2; DB 2; Length 44377;
Best Local Similarity 49.3%; Pred. No. 0.69;
Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 5 GGAGAAGCGCTACGACCTCGCGCGGTGGCGGTACAAAGGTCAACAAGAAGCTCGGCT 64
Db 20997 GGTGCGGGGCGCGGTCAACCCAGGACGGTCCACGACGGGCTCACCGCACCAGCGG 21056
QY 65 GAACACCGGCTCCCGATCAGCAGACCACTCTGACCGCAAGAGGAGCTGTCGCCACCAT 124
Db 21057 GCGCGCCGACGACGCGGTGATCGGTGACGCGGTGGCGGCGGTGACGCGCGCGCA 21116
QY 125 CGAGTACCTGTCGCGCTGTCACGAGGGCCACACGACGATGACGTCGCGGGGAGTCTGA 184
Db 21117 CGTGAGCGGTCGAGGCGGACGGCACCGCTCGGGGACCGCATCGAGGCGG 21176
QY 185 GGTGCGGTTGGAACCGACGACA 207
Db 21177 CGCGCTGATGGCCACCTACGGCA 21199

RESULT 8
US-08-804-198-1
; Sequence 1, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rostock, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAUL R. CANTRELL 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,198
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTRELL, PAUL R.
; REGISTRATION NUMBER: 36,470
; REFERENCE/DOCKET NUMBER: P9113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3885
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 350..14002
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31329..36071
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36155..41830
;
US-08-804-198-1

Query Match      18.4%; Score 38.2; DB 2; Length 44377;
Best Local Similarity 49.3%; Pred. No. 0.69;
Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 5 GGAGAAGCGCTACGACCTCGCGCGGTGGCGGTACAAAGTCAACAAAGAGTCTGGCGCT 64
Db 20997 GTGCGGGGCGAGCGGGTCAACACGAGGAGGTGCCAGCAACGGCTCACCGCACCCAGCGG 21056

QY 65 GAACACCCCGTCCCGATCAGCAGCAGCACCTCTGACCGAGAGGACGCTGTCGCCACCAT 124
Db 21057 GCCCGCCAGCAGGGGTGATCCGTGAGCGCTGGCCGACGCGGGGTGAGCGCCCGCA 21116

QY 125 CGAGTACTGTGTCGCCCTGACGAGGCGCACACACGATGACCTGCCGGCGGAGTCTGA 184
Db 21117 CGTGACGCGGTGAGGCGCACGCGCACCCGCGCACCCGCTCGGCGACCCGATCGAGGCGG 21176

QY 185 GGTCCCGTGAACCCGACGACA 207
Db 21177 CGCGCTGATGCCACCTACGGCA 21199

RESULT 9
US-09-513-783A-49
; Sequence 49, Application US/09513783A
; Patent No. 6416959
; GENERAL INFORMATION:
; APPLICANT: Giuliano, Kenneth A.
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-L1
; CURRENT APPLICATION NUMBER: US/09/513,783A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(717)
; OTHER INFORMATION: Description of Artificial Sequence: ECFP
US-09-513-783A-49

Query Match      18.3%; Score 38; DB 4; Length 717;
Best Local Similarity 49.0%; Pred. No. 0.53;
Matches 101; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGCTACGACCTCGCGGGTGGCGGTACAAAGTCAACAAGAAGCTCG 60
Db 392 TCAAGGAGGACGGCAACATCTCTGGGCGACAAAGCTGGAGTACAACTACATCAGCCACAAG 451

QY 61 GCCTGACACCGGTCCTCCCGATCAGCAGCACCACTCTGACCGAAGAGGAGCTGTCGCA 120
Db 452 TCTATATATACCCCGCAGCAGCAGCAAGCGCATCAAGGCCCACTTCAAGATCCGCCACA 511

QY 121 CCATCGAGTACCTGCTCGCGCTGACGAGGCGCCACACGATGACCTGCCGGCGGAG 180
Db 512 ACATCGAGGACGGCAGCTGTCAGCTCGCGGACCACTACGAGCAACACCCCATCGCGG 571

QY 181 TCGAGGTGCGCGTGGAAACCGACGAC 206
Db 572 ACGGCCCGCTGCTGCTGCCCGACAAC 597

RESULT 11
US-09-172-063-13
; Sequence 13, Application US/09172063
; Patent No. 6150176
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/172,063
; CURRENT FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 09/094,359
```

```
Db 452 TCTATATCAGCGCGGCAAGCAAGAAACGGCATCAAGGCCAACTTCAAGATCCGCCACA 511
QY 121 CCATCGAGTACCTGCTCGCGCTCAGCAGGCGCCACACGATGACCGTCCCGGGCGGAG 180
Db 512 ACATCGAGGACGGCAGCTGCTGCCGACCACTACGAGCAACACCCCATCGCGG 571
QY 181 TCGAGGTGCGCGTGGAAACCGACGAC 206
Db 572 ACGGCCCGCTGCTGCTGCCCGACAAC 597

RESULT 10
US-09-094-359-7
; Sequence 7, Application US/09094359
; Patent No. 6140132
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; FILE REFERENCE: 07257/067001
; CURRENT APPLICATION NUMBER: US/09/094,359
; CURRENT FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Aequorea victoria
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(717)
; OTHER INFORMATION: ECFP
US-09-094-359-7

Query Match      18.3%; Score 38; DB 3; Length 720;
Best Local Similarity 49.0%; Pred. No. 0.53;
Matches 101; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGCTACGACCTCGCGCGGTGGCGGTACAAAGTCAACAAGAAGCTCG 60
Db 392 TCAAGGAGGACGGCAACATCTCTGGGCGACAAAGCTGGAGTACAACTACATCAGCCACAAG 451

QY 61 GCCTGACACCGGTCCTCCCGATCAGCAGCACCACTCTGACCGAAGAGGAGCTGTCGCA 120
Db 452 TCTATATATACCCCGCAGCAGCAGCAAGCGCATCAAGGCCCACTTCAAGATCCGCCACA 511

QY 121 CCATCGAGTACCTGCTCGCGCTGACGAGGCGCCACACGATGACCTGCCGGCGGAG 180
Db 512 ACATCGAGGACGGCAGCTGTCAGCTCGCGGACCACTACGAGCAACACCCCATCGCGG 571

QY 181 TCGAGGTGCGCGTGGAAACCGACGAC 206
Db 572 ACGGCCCGCTGCTGCTGCCCGACAAC 597

RESULT 11
US-09-172-063-13
; Sequence 13, Application US/09172063
; Patent No. 6150176
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/172,063
; CURRENT FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 09/094,359
```


GT-ECFP
US-09-172-063-27

Query Match	18.38;	Score 38;	DB 3;	Length 972;
Best Local Similarity	49.06;	Pred. No. 0.54;		
Matches 101;	Conservative 0;	Mismatches 105;	Indels 0;	Gaps 0;
Qy 1	TCAAGGGAAGCGCTACGACCTCGCGGGGTGGCCGCTACAAGGTCAACAAGACGCTCG 60			
Db 644	TCAAGGACGACGGCAACATCTCTGGGCACAGCTGGAGTACAACACTACATCGGCCAACGG 703			
Qy 61	GCCTGAACACCGGTCCCGGATACGAGGAGCCACTCTGACCGAAGAGGACGCTGCTGGCCA 120			
Db 704	TCTATTATCACCCGCCACAAGCAGAAGAACGGCATCAAGGGCCCCACTTCAAGATCCGCCACA 763			
Qy 121	CCATCGAGTACCTTGTCGCCCTGCACGAGGGCCACACACGATGACCGTCCCGGGCGGAG 180			
Db 764	ACATCGAGGACGGCAGCGTGCAGCTGCCGCACCACTACCAGCAAAACCCCCCATCGCGG 823			
Qy 181	TCGAGGTCCGGTGGAAACCGACGAC 206			
Db 824	ACGGCCCGGTGCTGCTGCCCGACAAAC 849			

```

RESULT 15
US-09-513-783A-33
; Sequence 33, Application US/09513783A
; Patent No. 6416959
; GENERAL INFORMATION:
; APPLICANT: Giuliano, Kenneth A.
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-L1
; CURRENT APPLICATION NUMBER: US/09/513,783A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
; LENGTH: 1623
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1623)
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: VFP-NLS-CP3-multiple DEVD-CFP-Annexin II construct
US-09-513-783A-33

```

Query Match	18.38;	Score 38;	DB 4;	Length 1623;
Best Local Similarity	49.00;	Pred. No. 0.57;		
Matches 101;	Conservative 0;	Mismatches 105;	Indels 0;	Gaps 0;
QY 1	TCAAGGAGAAGCGCTACGACCTTCGCGGGGTGGSCCGCTACAAGGTCAACAAGAAGCTCG	60		
Db 1223	TCAAGGAGGACGGCAACATCTGGGGCACAAGCTGGAGTACAATTACATCAGGCCAACG	1282		
QY 61	GCGTGAACACGCGGTCCCGGATCACGACGACCACTCTGACCGAAGAGGACGCTGCTGGCCA	120		
Db 1283	TCTATATCACCCCGCACAAAGCAGAGAAGACGSCATCAAGGCCAACTTCAAGATCCGCCACA	1342		
QY 121	CCATCGAGTACCTGGTGGCGCTGCACGAGGGCCACACCCAGCATGCCGTCGCCGGCGGAG	180		
Db 1343	ACATCGAGACGGCAGCGTCAGCTGCCGACCACCTACCAGCAAAACACCCCATCGCGG	1402		
QY 181	TCGAGGTGCGGTGGAAACCGACGAC	206		
Db 1403	ACGGCCCGGTGCTGCTGCCCAAC	1428		

Search completed: November 12, 2002, 20:15:24
Job time : 81.8936 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 : Search time 136.222 Seconds
(without alignments)
3438.621 Million cell updates/sec

Title: US-09-697-123B-17

Perfect score: 208

Sequence: 1 tcaaggaagcgtacgac.....ccggtgaaaccgacacat 208

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. Is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	208	100.0	208	22	AAS05217
2	174.4	83.8	208	22	AAS05208
3	168	80.8	208	22	AAS05215
4	168	80.8	208	22	AAS05219
5	166.4	80.0	208	22	AAS05201
6	164.8	79.2	208	22	AAS05211
7	164.8	79.2	208	22	AAS05224
8	161.6	77.7	208	22	AAS05203
9	161.6	77.7	208	22	AAS05210

10	160	76.9	208	22	AAS05205	Mycobacterium tube
11	160	76.9	208	22	AAS05206	Mycobacterium terr
12	160	76.9	208	22	AAS05216	Mycobacterium bovi
13	160	76.9	970	17	AAT09676	Mycobacterium tube
14	160	76.9	3519	22	AAH51976	Mycobacterium tube
15	160	76.9	3534	22	AAH02079	Mycobacterium tube
16	160	76.9	3853	21	AAA74651	M. tuberculosis rp
17	160	76.9	3853	21	AAA89994	Mycobacterium afri
18	158.4	75.2	208	22	AAS05222	Mycobacterium gord
19	156.8	75.4	208	22	AAS05202	Mycobacterium haem
20	156.8	75.4	208	22	AAS05223	Mycobacterium szul
21	148	71.2	207	22	AAS05212	Mycobacterium gord
22	146.4	70.4	207	22	AAS05204	Mycobacterium intr
23	145.6	70.0	205	22	AAS05220	M. leprae rpoB gene
24	144	69.2	3447	14	AA051532	Mycobacterium flav
25	142.2	68.4	211	22	AA05218	Mycobacterium malm
26	140.8	67.7	214	22	AA05214	Mycobacterium absc
27	140.8	67.7	214	22	AA05221	Mycobacterium gast
28	138.2	66.4	223	22	AAS05213	Mycobacterium chel
29	134.4	64.6	214	22	AAS05207	Mycobacterium scro
30	128.6	61.8	223	22	AAS05209	C glutamicum codin
31	81.6	39.2	3495	22	AAH65512	C glutamicum codin
32	81.6	39.2	34980	22	AAH68525	Propionibacterium
33	69.2	33.3	27426	23	AAS59541	S. tendae nikkomyc
34	42.2	20.3	2918	21	AA243922	Streptomyces tende
35	42.2	20.3	2919	21	AA244486	M. capsulatus gene
36	41.6	20.0	2826	24	ABQ90502	M. capsulatus gene
37	41.6	20.0	49999	24	ABQ90984	Burkholderia cepac
38	41.4	19.9	1041	21	AA251708	Burkholderia cepac
39	41.4	19.9	1041	21	AA251709	Synthetic RT polyn
40	41.4	19.9	1680	24	ABL39997	Synthetic protease
41	41.4	19.9	1978	24	ABL39991	Synthetic Pol poly
42	41.4	19.9	3009	24	ABL40024	Synthetic Pol poly
43	41.4	19.9	3015	24	ABL39983	Synthetic Pol poly
44	41.2	19.8	2205	24	ABL50199	Human RapiA, c-Raf
45	41.2	19.8	2205	24	ABL50231	Human RapiA, c-Raf

ALIGNMENTS

RESULT 1
AAS05217
ID AAS05217 standard; DNA; 208 BP.
XX
AC AAS05217;
DT 07-SEP-2001 (first entry)
XX

XX
DE Mycobacterium celatum rpoB gene fragment.

XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.

XX
OS Mycobacterium celatum.

XX
PN WO200131061-A1.

XX
PD 03-MAY-2001.

XX
PF 27-OCT-2000; 2000WO-KR01223.

XX
PR 27-OCT-1999; 99KR-0046795.

XX
PA (ERUM-) ERUME BIOTECH CO LTD.

XX
PI Lee H, Park YK, Bai G, Kim S, Kim Y, Park HJ;

XX
DR WPI; 2001-300520/31.

XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism

```

XX PS Claim 1; Page 45; 50pp; English.
XX CC
XX CC The present sequence for Mycobacterium celatum rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX SQ Sequence 208 BP; 48 A; 71 C; 64 G; 25 T; 0 other;

Query Match 100.0%; Score 208; DB 22; Length 208;
Best Local Similarity 100.0%; Pred. No. 5.4e-38;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGGTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCG 60
DB 1 TCAAGGAGAAGCGGTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCG 60

QY 61 GCCTCAACACCGGTCGCCGCTGACGAGCACCCTCTGACCGAAGAGAGCGTCTCGCCA 120
DB 61 GCCTGAACACCGGTCGCCGCTGACGAGCACCCTCTGACCGAAGAGAGCGTCTCGCCA 120

QY 121 CCATCGAGTACCTGGTCCGCGCTGACGAGGGCCACACCGATGACCGTCCCGGGCGGAG 180
DB 121 CCATCGAGTACCTGGTCCGCGCTGACGAGGGCCACACCGATGACCGTCCCGGGCGGAG 180

QY 181 TCGAGGTCCGGTGGAAACCGACGACAT 208
DB 181 TCGAGGTCCGGTGGAAACCGACGACAT 208

RESULT 2
AAS05208
ID AAS05208 standard; DNA; 208 BP.
XX AC AAS05208;
XX DT 07-SEP-2001 (first entry)
XX DE Mycobacterium kansasii rpoB gene fragment.
XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX KW PCR-restriction fragment length polymorphism analysis; ds.
XX OS Mycobacterium kansasii.
XX PN WO200131061-A1.
XX PD 03-MAY-2001.
XX PF 27-OCT-2000; 2000WO-KR01223.
XX PR 27-OCT-1999; 99KR-0046795.
XX PA (ERUM-) ERUME BIOTECH CO LTD.
XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

```

```

XX WPI; 2001-300520/31.
XX DR
XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for
XX PT diagnosis and identification of many mycobacterial species by
XX PT restriction fragment length polymorphism
XX PS Claim 1; Page 42; 50pp; English.
XX CC
XX CC The present sequence for Mycobacterium kansasii rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX SQ Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;

Query Match 83.8%; Score 174.4; DB 22; Length 208;
Best Local Similarity 89.9%; Pred. No. 1.9e-30;
Matches 187; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGGTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCG 60
DB 1 TCAAGGAGAAGCGGTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCG 60

QY 61 GCCTGAACACCGGTCGCCGCTGACGAGCACCCTCTGACCGAAGAGAGCGTCTCGCCA 120
DB 61 GCCTGAACACCAATCATCCGATCACCACGACGCGTGACCGAAGAGAGCGTCTCGCCA 120

QY 121 CCATCGAGTACCTGGTCCGCGCTGACGAGGGCCACACCGATGACCGTCCCGGGCGGAG 180
DB 121 CCATCGAGTATCTGGTCCGCGCTGACGAGGGCCACCGATGACCGTCCCGGGCGGAG 180

QY 181 TCGAGGTCCGGTGGAAACCGGACGACAT 208
DB 181 TCGAGGTCCGGTGGAAACCGGACGACAT 208

RESULT 3
AAS05215
ID AAS05215 standard; DNA; 208 BP.
XX AC AAS05215;
XX DT 07-SEP-2001 (first entry)
XX DE Mycobacterium avium rpoB gene fragment.
XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX KW PCR-restriction fragment length polymorphism analysis; ds.
XX OS Mycobacterium avium.
XX PN WO200131061-A1.
XX PD 03-MAY-2001.
XX PF 27-OCT-2000; 2000WO-KR01223.

```

XX PR 27-OCT-1999; 99KR-0046795.
XX PA (ERUM-) ERUME BIOTECH CO LTD.
XX PI Lee H, Park YK, Bai G, Kim S, Kim Y, Park HJ;
XX DR WPI; 2001-300520/31.
XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for
XX PT diagnosis and identification of many mycobacterial species by
XX PT restriction fragment length polymorphism -
XX PS Claim 1; Page 44; 50pp; English.
XX CC The present sequence for Mycobacterium avium rpoB gene
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX CC various Mycobacterial species. These rpoB gene fragments can be used
XX CC in the diagnosis and identification of Mycobacterium species using a
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)
XX CC method. The method comprises obtaining a restriction fragment length
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX CC amplifying and digesting the DNA fragment from the microorganism to
XX CC be identified and comparing the RFLP patterns from the known rpoB gene
XX CC fragments with the unidentified fragment. The rpoB gene fragments
XX CC are useful to identify a wide range of Mycobacterium species, e.g. for
XX CC diagnosis or to obtain epidemiological and pathogenesis information for
XX CC selection of appropriate therapies, including M. tuberculosis, M. leprae
XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR
XX CC required), and can differentiate between many species in a single
XX CC experiment, including those difficult to distinguish by usual biochemical
XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX CC detecting specific Mycobacterial species.
XX SQ Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;
Query Match 80.8%; Score 168; DB 22; Length 208;
Best Local Similarity 88.0%; Pred. No. 5.1e-29;
Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTCGCGGGGTGGCGCTACAGGTCAACAGAGCTCG 60
Db 1 TCAAGGAGAGCGCTACGACCTCGCGGGGTGGCGCTACAGGTCAACAGAGCTCG 60
QY 61 GCCTGAACACCGCGTCCCGGATCAGCAGGACCTCTGACCGAGAGGACGTCGCGCCA 120
Db 61 GCCTGACGCGGTGAGCCGATCACCAGCTCGACGCTGACCGAGGAGAGCTCGCGCCA 120
QY 121 CCATCGAGTACCTGGTCCCGCTCGACGAGGGCCACACCGATGACCGTCCCGGGCGGAG 180
Db 121 CCATCGAGTACCTGGTCCCGCTCGACGAGGGTTCAGCCACGATGACCGTCCCGGGCGGCA 180
QY 181 TCGAGGTGCGGTGGAAACCGACGACAT 208
Db 181 TCGAGGTGCGGTGGAGACCGACGACAT 208
RESULT 4
AAS05219
ID AAS05219 standard; DNA; 208 BP.
XX AC AAS05219;
XX DT 07-SEP-2001 (first entry)
XX DE Mycobacterium fortuitum rpoB gene fragment.
XX DE Mycobacterium fortuitum.
XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX KW PCR-restriction fragment length polymorphism analysis; ds.
XX OS Mycobacterium fortuitum.

XX PN WO200131061-A1.
XX PD 03-MAY-2001.
XX PF 27-OCT-2000; 2000WO-KR01223.
XX PR 27-OCT-1999; 99KR-0046795.
XX PA (ERUM-) ERUME BIOTECH CO LTD.
XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX DR WPI; 2001-300520/31.
XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for
XX PT diagnosis and identification of many mycobacterial species by
XX PT restriction fragment length polymorphism -
XX PS Claim 1; Page 46; 50pp; English.
XX CC The present sequence for Mycobacterium fortuitum rpoB gene
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX CC various Mycobacterial species. These rpoB gene fragments can be used
XX CC in the diagnosis and identification of Mycobacterium species using a
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)
XX CC method. The method comprises obtaining a restriction fragment length
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX CC amplifying and digesting the DNA fragment from the microorganism to
XX CC be identified and comparing the RFLP patterns from the known rpoB gene
XX CC fragments with the unidentified fragment. The rpoB gene fragments
XX CC are useful to identify a wide range of Mycobacterium species, e.g. for
XX CC diagnosis or to obtain epidemiological and pathogenesis information for
XX CC selection of appropriate therapies, including M. tuberculosis, M. leprae
XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR
XX CC required), and can differentiate between many species in a single
XX CC experiment, including those difficult to distinguish by usual biochemical
XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX CC detecting specific Mycobacterial species.
XX SQ Sequence 208 BP; 43 A; 69 C; 69 G; 27 T; 0 other;
Query Match 80.8%; Score 168; DB 22; Length 208;
Best Local Similarity 88.0%; Pred. No. 5.1e-29;
Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTCGCGGGGTGGCGCTACAGGTCAACAGAGCTCG 60
Db 1 TCAAGGAGAGCGCTACGACCTCGCGGGGTGGCGCTACAGGTCAACAGAGCTCG 60
QY 61 GCCTGAACACCGCGTCCCGGATCAGCAGGACCTCTGACCGAGAGGACGTCGCGCCA 120
Db 61 GCCTGACGCGGTGAGCCGATCACCAGCTCGACGCTGACCGAGGAGAGCTCGCGCCA 120
QY 121 CCATCGAGTACCTGGTCCCGCTCGACGAGGGCCACACCGATGACCGTCCCGGGCGGAG 180
Db 121 CCATCGAGTACCTGGTCCCGCTCGACGAGGGTTCAGCCACGATGACCGTCCCGGGCGGAG 180
QY 181 TCGAGGTGCGGTGGAAACCGACGACAT 208
Db 181 TCGAGGTGCGGTGGAGACCGACGACAT 208
RESULT 5
AAS05201
ID AAS05201 standard; DNA; 208 BP.
XX AC AAS05201;
XX DT 07-SEP-2001 (first entry)
XX XX

DE Mycobacterium gordonae type I rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium gordonae type I.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism

PS Claim 1; Page 40; 50pp; English.

CC The present sequence for Mycobacterium gordonae type I rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 46 A; 71 C; 65 G; 26 T; 0 other;

Query Match 80.0%; Score 166.4; DB 22; Length 208;

Best Local Similarity 87.5%; Pred. No. 1.2e-28;

Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTAGGACCTCGCGCGGTGGCGGTACAAAGTCAACAGAAGCTCG 60

DB 1 TCAAGGAGAGCGCTAGGACCTCGCGCGGTGGCGGTACAAAGTCAACAGAAGCTCG 60

QY 61 GCCTGAACACCGCGTCCCGGATCACCACGACCACTCTGACCGAAGAGACGTCGCGCA 120

DB 61 GCCTGAACACCGCGTCCCGGATCACCACGCTCCACCGTACCGAGGAAGACGTCGCGCA 120

QY 121 CCATCGAGTACCTGGTCCGCTGCAGAGGGCCACACGATGACCGTCCCGGCGCGGAG 180

DB 121 CCATCGAGTACCTGGTCCGCTGCAGAGGGCCACACGATGACCGTCCCGGCGCGGAG 180

QY 181 TCGAGGTGCGGTTGGAACCGAGACAT 208

DB 181 CCGAGGTGCGGTTGAGCCGAGACAT 208

RESULT 6
AAS05211

ID

XX AAS05211 standard; DNA; 208 BP.

XX AAS05211;

XX 07-SEP-2001 (first entry)

XX Mycobacterium marinum rpoB gene fragment.

DE Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium marinum.

XX WO200131061-A1.

PN 03-MAY-2001.

PD 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI; 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism

PS Claim 1; Page 43; 50pp; English.

CC The present sequence for Mycobacterium marinum rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 45 A; 59 C; 68 G; 26 T; 0 other;

Query Match 79.2%; Score 164.8; DB 22; Length 208;

Best Local Similarity 87.0%; Pred. No. 2.6e-28;

Matches 181; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTAGGACCTCGCGCGGTGGCGGTACAAAGTCAACAGAAGCTCG 60

DB 1 TCAAGGAGAGCGCTAGGACCTCGCGCGGTGGCGGTACAAAGTCAACAGAAGCTCG 60

QY 61 GCCTGAACACCGCGTCCCGGATCACCACGACCACTCTGACCGAAGAGACGTCGCGCA 120

DB 61 GCCTGAACACCGCGTCCCGGATCACCACGCTCCACCGTACCGAGGAAGACGTCGCGCA 120

QY 121 CCATCGAGTACCTGGTCCGCTGCAGAGGGCCACACGATGACCGTCCCGGCGCGGAG 180

DB 121 CCATCGAGTACCTGGTCCGCTGCAGAGGGCCACACGATGACCGTCCCGGCGCGGAG 180

QY 181 TCGAGGTGCGGTTGGAACCGAGACAT 208

```

Db 181 TCGAGGTCCCGTGCAGACCGACGACAT 208
|||||
RESULT 7
AAS05224
ID AAS05224 standard; DNA: 208 BP.
XX
AC AAS05224;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium xenopi rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium xenopi.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI: 2001-300520/31.
XX
DR New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX Claim 1; Page 47; 50pp; English.
XX
CC The present sequence for Mycobacterium xenopi rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC experiment), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 51 A; 65 C; 67 G; 25 T; 0 other;

Query Match 79.2%; Score 164.8; DB 22; Length 208;
Best Local Similarity 87.0%; Pred. No. 2.6e-28;
Matches 181; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAGAAAGCTCG 60
Db 1 TCAAGGAGAACGGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAGAAAGCTCG 60
QY 61 GCCTGAACACCGCTCCCGGATCAGGACGACCTCTGACCGAGAGAGCGTGTGCGCCA 120
Db 61 GCGTGAACACCGAGTCCGCCACCCACCGACCTTGACCGAGAGAGGAGCGTGTGCGCCA 120

```

```

QY 121 CCATCGAGTACCTGGTCCGCTGCACGAGGGCCACACGACGATGACCGTCCCGGGCGGAG 180
|||||
Db 121 CCATCGAATACCTGGTGGCTTGCAGGGGCGACCCACGATGAAGGTCCCGGTGGCG 180
|||||
QY 181 TCGAGGTCCCGGTGGAACCGACGACAT 208
|||||
Db 181 TCGAGGTCCCGGTGGAACCGACGACAT 208
|||||

```

RESULT 8

```

AAS05203
ID AAS05203 standard; DNA: 208 BP.
XX
AC AAS05203;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium gordonae type III rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium gordonae type III.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI: 2001-300520/31.
XX
DR New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX Claim 1; Page 41; 50pp; English.
XX
CC The present sequence for Mycobacterium gordonae type III rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC experiment), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 46 A; 69 C; 63 G; 30 T; 0 other;

```

```

Query Match 77.7%; Score 161.6; DB 22; Length 208;
Best Local Similarity 86.1%; Pred. No. 1.4e-27;
Matches 179; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAGAAAGCTCG 60

```

Db 1 TCAAGGAGAGCGGTACGACCTGCGCCGTGTCGCGCGCTACAAGGTCAACAAGAGCTCG 60
QY 61 GCCTGAACACCGCGTCCCGGATCAGGACGACCTCTGACCGAAGAGAGGTGCTCGCCCA 120
Db 61 GCCTGACAGTCCGGGATCCGATCACCAAGTCCACGCTGACCGAAGAAGAGAGTCTCGCCCA 120
QY 121 CCATCGACTACTGTCGCGCTGCACGAGGCGCCACACACGATGACCGTCCCGGGCGGAG 180
Db 121 CCATCGAGTACTGTCGCTGTCACGAGGTCAGCAGATGACCGTTCGCGGCGGCA 180
QY 181 TCGAGGTCCCGTGAACACGACGACAT 208
Db 181 CCGAGTTCCGGTGAGACGACGACAT 208
RESULT 9
AAS05210
ID AAS05210 standard; DNA; 208 BP.
AC AAS05210;
XX
XX 07-SEP-2001 (first entry)
DE Mycobacterium ulcerans rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
XX Mycobacterium ulcerans.
OS
XX WO200131061-A1.
PN
XX
XX 03-MAY-2001.
PD
XX
XX 27-OCT-2000; 2000WO-KR01223.
PF
XX
XX 27-OCT-1999; 99KR-0046795.
PR
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
PA
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
PI
XX
XX WPI; 2001-300520/31.
DR
XX
XX The present sequence for Mycobacterium ulcerans rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
XX Sequence 208 BP: 45 A; 68 C; 67 G; 28 T; 0 other;

Query Match 77.7%; Score 161.6; DB 22; Length 208;
Best Local Similarity 86.1%; Pred. No. 1.4e-27;
Matches 179; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGGTACGACCTGCGCGGTGCGCGCTACAAGGTCAACAAGAGCTCG 60
Db 1 TCAAGGAGAGCGGTACGACCTGCGCTGCGGTGCGGTACAAGGTCAACAAGAGCTCG 60
QY 61 GCCTGAACACCGCGTCCCGGATCAGCAGCACCCTCTGACCGAAGAGAGGTGCTCGCCCA 120
Db 61 GCCTGAACCGCGGCCAGCCCATCACAGCTCGACGCTGACCGAAGAAGAGTCTCGTCGCCCA 120
QY 121 CCATCGAGTACTGTCGCGCTGCACGAGGCGCCACACACGATGACCGTCCCGGGCGGAG 180
Db 121 CCATCGAATACCTGTCGCTTGCACGAGGCGCAGACCGCGATGACCGTCCCGGGCGGTG 180
QY 181 TCGAGGTCCCGTGAACACGACGACAT 208
Db 181 TCGAGGTCCCGTGCAGACCGAGACAT 208
RESULT 10
AAS05205
ID AAS05205 standard; DNA; 208 BP.
XX
AC AAS05205;
XX
XX 07-SEP-2001 (first entry)
DE Mycobacterium tuberculosis rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
XX Mycobacterium tuberculosis.
OS
XX WO200131061-A1.
PN
XX
XX 03-MAY-2001.
PD
XX
XX 27-OCT-2000; 2000WO-KR01223.
PF
XX
XX 27-OCT-1999; 99KR-0046795.
PR
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
PA
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
PI
XX
XX WPI; 2001-300520/31.
DR
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
XX Disclosure; Page 41; 50pp; English.
XX
XX The present sequence for Mycobacterium tuberculosis rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
XX Sequence 208 BP: 45 A; 68 C; 67 G; 28 T; 0 other;

CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
Query Match 76.9%; Score 160; DB 22; Length 208;
Best Local Similarity 85.6%; Pred. No. 3.2e-27;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1 TCAAGGAGAACGCTACGACCTCGCGGGTGGCGCTACAGGTCAACAAGAAGCTCG 60
DB 1 TCAAGGAGAACGCTACGACCTGCGCGCTATAGGTCAACAAGAAGCTCG 60
QY 61 GCCTGAACACCGCTCCCGATCAGACGACCACTCTGACCGAAGAGAGCTCGTCCGCA 120
DB 61 GCCTGCATGTGCGGAGGCCATCATCGTCTGACGCTGACCGAAGAAGAGCTCGTGGCCA 120
QY 121 CCATCGAGTACCTGCTCGGCTGCACGAGGCCACACACGATGACCGTCCCGGGCGAG 180
DB 121 CCATCGAATATCTGCTCGCTTGCACGAGGGTCAGACCATGACCGTTCCGGGGCGG 180
QY 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208
DB 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208
RESULT 11
AAS05206
ID AAS05206 standard; DNA; 208 BP.
AC AAS05206;
XX
XX 07-SEP-2001 (first entry)
XX Mycobacterium terrae rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
XX Mycobacterium terrae.
XX
XX WO200131061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI: 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism -
XX
XX Claim 1; Page 42; 50pp; English.
XX
XX The present sequence for Mycobacterium terrae rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments

CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
Query Match 76.9%; Score 160; DB 22; Length 208;
Best Local Similarity 85.6%; Pred. No. 3.2e-27;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1 TCAAGGAGAACGCTACGACCTCGCGGGTGGCGCTACAGGTCAACAAGAAGCTCG 60
DB 1 TCAAGGAGAACGCTACGACCTGCGCGCTATAGGTCAACAAGAAGCTCG 60
QY 61 GCCTGAACACCGCTCCCGATCAGACGACCACTCTGACCGAAGAGAGCTCGTCCGCA 120
DB 61 GCCTGCATGTGCGGAGGCCATCATCGTCTGACGCTGACCGAAGAAGAGCTCGTGGCCA 120
QY 121 CCATCGAGTACCTGCTCGGCTGCACGAGGCCACACACGATGACCGTCCCGGGCGAG 180
DB 121 CCATCGAATATCTGCTCGCTTGCACGAGGGTCAGACCATGACCGTTCCGGGGCGG 180
QY 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208
DB 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208
RESULT 12
AAS05216
ID AAS05216 standard; DNA; 208 BP.
XX
XX AAS05216;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium bovis rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
XX Mycobacterium bovis.
XX
XX WO200131061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI: 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism -
XX
XX Claim 1; Page 45; 50pp; English.
XX
XX The present sequence for Mycobacterium bovis rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a

CC susceptibility, particularly to rifamycin. The method can provide
CC often greater than 9% sensitivity and 100% specificity. The
CC biological sample is a fluid or tissue sample from a human.
XX
SQ Sequence 970 BP; 182 A; 302 C; 330 G; 156 T; 0 other;
Query Match 76.9%; Score 160; DB 17; Length 970;
Best Local Similarity 85.6%; Pred. No. 3.2e-27;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAAGCGCTACGACCTCCCGCGGTGGCGGTGCTACAAAGTCAACAAGAGCTCG 60
Db 26 TCAAGGAGAAGCGCTACGACCTCCCGCGGTGGCGGTGCTACAAAGTCAACAAGAGCTCG 85
Qy 61 GCCTGAACACCGCGTCCCGCATCAGCAGCACCCTCTGACCGAAGAGGACGCTCGGCCA 120
Db 86 GCCTGCATGTGCGGGAGCCCATCAGCTGCTGACGCTGACCGAAGAGAGCTCGTGCCCA 145
Qy 121 CCATCGAGTACCTGTCGCGCTCCACGAGGGCCACACCATGATGACCGTCCCGGGCGGAG 180
Db 146 CCATCGAATATCTGTCGCGCTTGCAGAGGTCAGACCATGATGACCGTCCCGGGCGCG 205
Qy 181 TCGAGGTCCCGTGGAAACCGACAT 208
Db 206 TCGAGGTCCCGTGGAAACCGACAT 233
RESULT 14
ID AAH51976 standard; DNA; 3519 BP.
XX
AC AAH51976;
XX
DT 04-SEP-2001 (first entry)
XX
DE Mycobacterium tuberculosis potential drug target gene SEQ ID 30.
XX
KW Drug target; growth; organism viability; characterisation; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200135317-A1.
XX
PD 17-MAY-2001.
XX
PF 13-NOV-2000; 2000WO-US31152.
XX
PR 12-NOV-1999; 99US-0165086.
PR 12-NOV-1999; 99US-0165124.
PR 01-FEB-2000; 2000US-0179531.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Eisenberg D, Rotstein SH, Marcotte EM;
XX
XX WPI: 2001-329193/34.
DR P-PSDB; AAG81125.
XX
XX Identifying nucleotide or polypeptide sequence for use as drug target,
PT involves providing algorithm that analyzes a functional relationship
PT between nucleotide or polypeptide sequences, and comparing the
PT sequences.
XX
XX Disclosure: Page 68-69; 207pp; English.
XX
XX This invention relates to a method for identifying a nucleotide or
CC polypeptide sequence that may be a drug target, or essential for growth
CC or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092
CC represent DNA encoding proteins AAG81096 - AAG81241, Mycobacterium
CC tuberculosis proteins which are potential drug targets. The DNA and
CC protein sequences are used to illustrate the method of the invention. The
CC method involves providing an unknown nucleotide or polypeptide sequences,
CC and comparing it to a number of sequences along with at least one

CC algorithm capable of analysing a functional relationship between
CC nucleotide and polypeptide sequences. The method is useful for
CC characterising the function of nucleic acids and polypeptides that may be
CC useful as a target for a drug or essential for the growth or viability of
CC an organism.
XX
SQ Sequence 3519 BP; 675 A; 1081 C; 1183 G; 580 T; 0 other;
Query Match 76.9%; Score 160; DB 22; Length 3519;
Best Local Similarity 85.6%; Pred. No. 3.2e-27;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAAGCGCTACGACCTCCCGCGGTGGCGGTGCTACAAAGTCAACAAGAGCTCG 60
Db 884 TCAAGGAGAAGCGCTACGACCTCCCGCGGTGGCGGTGCTACAAAGTCAACAAGAGCTCG 943
Qy 61 GCCTGAACACCGCGTCCCGCATCAGCAGCACCCTCTGACCGAAGAGGACGCTCGGCCA 120
Db 944 GCCTGCATGTGCGGGAGCCCATCAGCTGCTGACGCTGACCGAAGAGAGCTCGTGCCCA 1003
Qy 121 CCATCGAGTACCTGTCGCGCTCCACGAGGGCCACACCATGATGACCGTCCCGGGCGGAG 180
Db 1004 CCATCGAATATCTGTCGCGCTTGCAGAGGTCAGACCATGATGACCGTCCCGGGCGCG 1063
Qy 181 TCGAGGTCCCGTGGAAACCGACAT 208
Db 1064 TCGAGGTCCCGTGGAAACCGACAT 1091
RESULT 15
ID AAH02079 standard; DNA; 3534 BP.
XX
AC AAH02079;
XX
DT 24-JUL-2001 (first entry)
XX
DE Mycobacterium tuberculosis nucleotide sequence SEQ ID NO:2072.
XX
KW Species specific; genus specific; family specific; probe; detection;
KW identification; algal; archaeal; bacterial; fungal; parasitical;
KW microorganism; diagnosis; translation elongation factor tu; toxin;
KW translation elongation factor G; RecA recombinase; resistance;
KW catalytic subunit of proton-translocating ATPase; antimicrobial;
KW vaccine; primer; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200123604-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-CA01150.
XX
PR 28-SEP-1999; 99CA-2283458.
PR 19-MAY-2000; 2000CA-2307010.
XX
XX (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX
XX Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
PI Picard FJ, Roy PH;
XX
XX WPI: 2001-245006/25.
XX
XX Nucleic acid sequences are used to generate universal probes and
PT primers which can be used to identify and detect the presence of algal,
PT archaeal, bacterial, fungal and parasitical species in a test sample.
XX
XX Disclosure: Page 1478-1479; 1580pp; English.
XX
XX The present invention describes a method for generating a repository of
CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
CC and/or primers are derived. The method comprises amplifying the nucleic

CC acids of determined algal, archaeal, bacterial, fungal and parasitological
CC species with a combination of defined primer pairs. The method can be
CC used for producing probes and/or primers for detecting one or more
CC related microorganisms e.g. algae, archaea, bacteria, fungi and
CC parasites, for universal detection and for specific and ubiquitous
CC detection and identification of an algal, archaeal, bacterial, fungal
CC and parasitological species, genus, family and group. A nucleic acid (1)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexA nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (1) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
CC which are given in the exemplification of the present invention.
XX
SQ Sequence 3534 BP: 679 A; 1081 C; 1188 G; 586 T; 0 other;

Query Match 76.9%; Score 160; DB 22; Length 3534;
Best Local Similarity 85.6%; Pred. No. 3.2e-27;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCG 60
DB 902 TCAAGGAGAGCGCTACGACCTCGCGGGTGGCGGCTACAAGGTCAACAAGAAGCTCG 961
QY 61 GCCTGAACACCGCTCCCGGATCACGACGACCACTCTGACCGAAGAGGAGCTCGTGGCCA 120
DB 962 GCCTGCACTGTCGGCGAGCCCATACGCTCGACGCTGACCGAAGAGGAGCTCGTGGCCA 1021
QY 121 CCATCGAGTACCTGGTCCGCTGCGAGGGCCACACACGATGACCGTCCCGGGCGGAG 180
DB 1022 CCATCGAATATCTGGTTCGCTTGCAGAGGGTTCAGACCACGATGACCGTCCCGGGCGG 1081
QY 181 TCGAGGTGCGGTGGAAACCGACGACAT 208
DB 1082 TCGAGGTGCGGTGGAAACCGACGACAT 1109

Search completed: November 12, 2002, 16:50:16
Job time : 140.222 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 ; Search time 636.023 Seconds
(without alignments)
9517.553 Million cell updates/sec

Title: US-09-697-123B-17

Perfect score: 208

Sequence: 1 tcaaggagaagcgtacgac.....ccggtggaacacgacacat 208

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rtd.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	160	76.9	610	1	MTU318818	AJ318818 Mycobacte
2	160	76.9	610	1	MTU318819	AJ318819 Mycobacte
3	160	76.9	616	1	MTU318813	AJ318813 Mycobacte
4	160	76.9	618	1	MTU318815	AJ318815 Mycobacte
5	160	76.9	618	1	MTU318817	AJ318817 Mycobacte
6	160	76.9	633	1	MTU318814	AJ318814 Mycobacte
7	160	76.9	637	1	MTU318816	AJ318816 Mycobacte
8	160	76.9	639	1	MTU318821	AJ318821 Mycobacte
9	160	76.9	970	6	I50706	I50706 Sequence 1
10	160	76.9	3534	6	AX111339	AX111339 Sequence
11	160	76.9	3853	1	MTU12205	MTU12205 Mycobacteri
12	160	76.9	5084	1	MSGRPOB	L27989 Mycobacteri
13	160	76.9	19352	1	AE006964	AE006964 Mycobacte
14	160	76.9	19770	1	MTC1376	295972 Mycobacteri
15	152	73.1	3752	1	MSU24494	U24494 Mycobacteri
16	144	69.2	3447	6	AR067447	AR067447 Sequence
17	144	69.2	37617	1	MLB1790G	Z14314 M.leprae ge
18	144	69.2	348950	1	MLEPRTN7	AL583923 Mycobacte
19	117.2	56.3	3941	1	AF242549	AF242549 Amycolato
20	83.6	40.2	32923	1	SCD82	AL160431 Streptomy
21	81.6	39.2	3495	6	AX120631	AX120631 Sequence
22	81.6	39.2	328050	1	AP005275	AP005275 Corynebac
23	81.6	39.2	349980	6	AX127144	AX127144 Sequence
24	49	23.6	18471	1	SCBAC17F8	AL596030 Streptomy
25	47.8	23.0	1416	1	SRRECAGEN	X94233 S.rimosus r
26	47	22.6	195859	14	AF281817	AF281817 Tupai a he
27	46	22.1	33001	1	SCC123	AL136518 Streptomy
28	45.6	21.9	7954	1	TAO19223	Y19223 Thermus aqu
29	45.2	21.7	4575	1	AB083790	AB083790 Thermus t
30	44.8	21.5	204050	1	AL646073	AL646073 Ralstonia
31	43.6	21.0	183046	8	AC087599	AC087599 Oryza sat
32	43.4	20.9	11103	1	AE004774	AE004774 Pseudomon
33	43.4	20.9	14210	1	AE012929	AE012929 Chloroblu
34	43.2	20.8	20054	1	SCCB12	AL391588 Streptomy
35	42.8	20.6	2966	8	AF247649	AF247649 Conlolyr
36	42.8	20.6	23087	1	SCD40A	AL161691 Streptomy
37	42.8	20.6	30753	1	SCC54	AL035591 Streptomy
38	42.6	20.5	4333	1	AF012127	AF012127 Thlobacil
39	42.6	20.5	4556	1	TFSCSAB	X54073 T.flavus sc
40	42.4	20.4	12112	1	AE005722	AE005722 Caulobact
41	42.4	20.4	30000	6	AX250262	AX250262 Sequence
42	42.2	20.3	3291	1	MSRECAGEN	X94208 Mycobacteri
43	42.2	20.3	7870	1	STE18574	Y18574 Streptomyce
44	42	20.2	204050	1	AL646070	AL646070 Ralstonia
45	41.8	20.1	1380	1	PSEARGFP	M19939 P.aeruginos

ALIGNMENTS

RESULT 1
MTU318818

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

MTU318818 610 bp DNA circular BCT 09-AUG-2002
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1415-97.

AJ318818

RNA polymerase beta subunit; rpoB gene.

GR:22208412

Mycobacterium tuberculosis

Mycobacterium tuberculosis

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium

tuberculosis complex.

1

Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 610)

AUTHORS Herrera,L.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES

source

1. .610

/organism="Mycobacterium tuberculosis"

/isolate="1415-97"

/db_xref="taxon:1773"

gene

1. .610

/gene="rpoB"

CDS

<1. .>610

/gene="rpoB"

/codon_start=1

/transl_table=11

/product="RNA polymerase beta subunit"

/protein_id="CAC87035.1"

/db_xref="GI:22208413"

/translation="LDIVYKLRPGEPPTKESAOPTLLENLFKEKRYDLARVGRYKVK KLGLHVGEPITSLTEEDVATIEYLVRLHEGQTTMTVPGGVEVPVETDDIDHFGNR RLRTVGEIQNIQIRVGMSEMRVVRMTQDVEAITPQTLINIRPVVAAIKEFFGTS QPSQFMQNNPLSLGTHKRRLSALGPGLSRERAGLEVRDHP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 76.9%; Score 160; DB 1; Length 610;

Best Local Similarity 85.6%; Pred. No. 1e-17;

Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGGCTACAAGGTCAACAAGAGCTCG 60

|||||

Db 80 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGGCTACAAGGTCAACAAGAGCTCG 139

QY 61 GCCTGACACCGCTCCCGGATCAGCAGGACCACTCTGACCGAAGAGAGCTCGTCGCCA 120

|||||

Db 140 GGCTGCATGTCGGCGAGCCCATCAGCTGTCGACGTCACCGAAGAGAGCTCGTCGCCA 199

QY 121 CCATCGAGTACCTGTCGCGCTTCGCGCTTGCACGAGGCTCAGCAGATGACCGTCCCGGCGGAG 180

|||||

Db 200 CCATCGAATATCTGTCGCTTGCACGAGGCTCAGCAGATGACCGTCCCGGCGGCG 259

QY 181 TCGAGGTCCCGGTGGAACCGGACGACAT 208

|||||

Db 260 TCGAGGTCCCGGTGGAACCGGACGACAT 287

RESULT 2

MTU318819

LOCUS

DEFINITION

610 bp DNA circular BCT 09-AUG-2002

Myobacterium tuberculosis partial rpoB gene for RNA polymerase beta subunit, isolate 1417-97.

ACCESSION

AJ318819

VERSION

GI:22208414

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

Myobacterium tuberculosis.

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

1

REFERENCE

AUTHORS

Herrera,L., Jimenez,M.S. and Saez,J.A.

TITLE

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

Unpublished

REFERENCE

2 (bases 1 to 610)

AUTHORS

Herrera,L.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES

source

1. .610

/organism="Mycobacterium tuberculosis"

/isolate="1417-97"

/db_xref="taxon:1773"

gene

1. .610

/gene="rpoB"

CDS

<1. .>610

/gene="rpoB"

/codon_start=1

/transl_table=11

/product="RNA polymerase beta subunit"

/protein_id="CAC87036.1"

/db_xref="GI:22208415"

/translation="LDIYKLRPGEPPTKESAOPTLLENLFKEKRYDLARVGRYKVK KLGLHVGEPITSLTEEDVATIEYLVRLHEGQTTMTVPGGVEVPVETDDIDHFGNR RLRTVGEIQNIQIRVGMSEMRVVRMTQDVEAITPQTLINIRPVVAAIKEFFGTS QPSQFMQNNPLSLGTHKRRLSALGPGLSRERAGLEVRDHP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 76.9%; Score 160; DB 1; Length 610;

Best Local Similarity 85.6%; Pred. No. 1e-17;

Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGGCTACAAGGTCAACAAGAGCTCG 60

|||||

Db 80 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGGCTACAAGGTCAACAAGAGCTCG 139

QY 61 GCCTGAAACCGCGTCCCGGATCAGCAGGACCACTCTGACCGAAGAGAGCTCGTCGCCA 120

|||||

Db 140 GGCTGCATGTCGGCGAGCCCATCAGCTGTCGACGTCACCGAAGAGAGCTCGTCGCCA 199

QY 121 CCATCGAGTACCTGTCGCGCTTCGCGCTTGCACGAGGCTCAGCAGATGACCGTCCCGGCGGAG 180

|||||

Db 200 CCATCGAATATCTGTCGCTTGCACGAGGCTCAGCAGATGACCGTTCGCGGCGGCG 259

QY 181 TCGAGGTCCCGGTGGAACCGGACGACAT 208

|||||

Db 260 TCGAGGTCCCGGTGGAACCGGACGACAT 287

RESULT 3

MTU318813

LOCUS

DEFINITION

616 bp DNA circular BCT 09-AUG-2002

Myobacterium tuberculosis partial rpoB gene for RNA polymerase beta subunit, isolate 1763-97.

ACCESSION

AJ318813

VERSION

GI:22208402

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

Myobacterium tuberculosis.

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

1

REFERENCE

AUTHORS

Herrera,L., Jimenez,M.S. and Saez,J.A.

TITLE

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

Unpublished

REFERENCE

2 (bases 1 to 616)

AUTHORS

Herrera,L.

TITLE

Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES

source

1. .616

/translation="TDEALLDIYRKLRPGEPTTKESAQTLLLENLFFKRYDLARVGR
YKVNKKLGLHVGEPITSTLDEEDVATIEYVRLHREGOTTMTVPVGGVEVPVETDDID
YFGNRRRLRTVGLIQNIRVGMRSRMVRERMTTODVEAITPOTLINIRPVVAIKE
FFGTSSQSPFMQGNPLSLGTHKRRLLSALGPGGLSRERAGLEVRDV"
BASE COUNT 124 a 192 c 207 g 95 t
ORIGIN

Query Match 76.9%; Score 160; DB 1; Length 618;
Best Local Similarity 85.6%; Pred. No. 1e-17;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAAGTCAACAAGAGCTCG 60
Db 95 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAAGTCAACAAGAGCTCG 154
QY 61 GCCTGAACACCGCGTCCCGATCAGCAGCACCTCTGACCGAAGAGGACGTCGTCGCCA 120
Db 155 GGCTGCATGTCGGGAGGCCATCAGCTGTCGAGCGCTGACCGAAGAGAGCGTCGTGCCA 214
QY 121 CCATCGAGTACTGTGTCGCTGCAGGAGGCCACACACATGACCGTCCCGGGCGGAG 180
Db 215 CCATCGAATATCTGTCGCTGTCAGGAGGTCAGACACATGACCGTTCCTGGCGGCG 274
QY 181 TCGAGTCCCGGTGGAACCGACGACAT 208
Db 275 TCGAGTCCCGGTGGAACCGACGACAT 302

RESULT 6
MTU318814
LOCUS
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
633 bp DNA circular BCT 09-AUG-2002
beta subunit, isolate 1058-97.
AJ318814
ACCESSION
VERSION AJ318814.1 GI:22208404
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE 1
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 633)
AUTHORS Herrera, L.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN

FEATURES
source
1..633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
gene
1..633
/gene="rpoB"
CDS
<1..>633
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87031.1"
/db_xref="GI:22208405"

translation="DNTVGTDEALLDIYRKLRPGEPTTKESAQTLLLENLFFKRYDL
ARVGRYKVNKKLGLHVGEPITSTLDEEDVATIEYVRLHREGOTTMTVPVGGVEVPVE
TDDIDHFNRRRLRTVGLIQNIRVGMRSRMVRERMTTODVEAITPOTLINIRPVV
AAIKFEFGTSQLSQFMDQNNPLSLGTYRRLSALGPGGLSRERAGLEVRDV"
BASE COUNT 129 a 195 c 210 g 99 t
ORIGIN

Query Match 76.9%; Score 160; DB 1; Length 633;
Best Local Similarity 85.6%; Pred. No. 1e-17;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAAGTCAACAAGAGCTCG 60
Db 110 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAAGTCAACAAGAGCTCG 169
QY 61 GCCTGAACACCGCGTCCCGATCAGCAGCACCTCTGACCGAAGAGGACGTCGTCGCCA 120
Db 170 GGCTGCATGTCGGGAGGCCATCAGCTGTCGAGCGCTGACCGAAGAGAGCGTCGTGCCA 229
QY 121 CCATCGAGTACTGTGTCGCTGCAGGAGGCCACACATGACCGTCCCGGGCGGAG 180
Db 230 CCATCGAATATCTGTCGCTGTCAGGAGGTCAGACACATGACCGTTCCTGGCGGCG 289
QY 181 TCGAGTCCCGGTGGAACCGACGACAT 208
Db 290 TCGAGTCCCGGTGGAACCGACGACAT 317

RESULT 7
MTU318816
LOCUS
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
637 bp DNA circular BCT 09-AUG-2002
beta subunit, isolate 1255-98.
AJ318816
ACCESSION
VERSION AJ318816.1 GI:22208408
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE 1
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 637)
AUTHORS Herrera, L.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN

FEATURES
source
1..637
/organism="Mycobacterium tuberculosis"
/isolate="1255-98"
/db_xref="taxon:1773"
gene
1..637
/gene="rpoB"
CDS
<1..>637
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87033.1"
/db_xref="GI:22208409"
/translation="VGTDEALLDIYRKLRPGEPTTKESAQTLLLENLFFKRYDLARV
GRYKVNKKLGLHVGEPITSTLDEEDVATIEYVRLHREGOTTMTVPVGGVEVPVETDD
IDHFNRRRLRTVGLIQNIRVGMRSRMVRERMTTODVEAITPOTLINIRPVVAI
KEFGTSQLSQFMDQNNPLSLGTHKRRMFALGPGGLSRERAGLEVRDHPHSH"
BASE COUNT 128 a 198 c 210 g 101 t
ORIGIN

Query Match 76.9%; Score 160; DB 1; Length 637;
Best Local Similarity 85.6%; Pred. No. 1e-17;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTCGCGCGGTGGCGCTACAAAGTCAACAAGAGCTCG 60

|||||TCAAGGAGAACGCTACGACCTGCGCGCGTCTGCGTCTATAGGTCAACAAGAGCTCG 160
Db 101
|||||GCCTGAACACCGCGTCCCGGATCAGACACACACTCTGACCGAAGAGAGCTCGTCGCCCA 120
Qy 61
|||||GGCTGCATGTGCGCGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGAGCTCGTGCCCA 220
Db 161
|||||CCATCGAGTACCTGCTCGCGCTGCACGAGGCGCACACACGATGACCGTCCCGGCGGAG 180
Qy 121
|||||CCATCGAATATCTGCTCGCGTTCGACGAGGTCAGACCGATGACCGTTCCTCCGGCGGCG 280
Db 221
|||||TCGAGGTGCGCGGTGGAACCGAGACAT 208
Qy 181
|||||TCGAGGTGCGCGGTGGAACCGAGACAT 308
Db 281
RESULT 8
LOCUS MTU318821 639 bp DNA circular BCT 09-AUG-2002
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1071-98.
ACCESSION AJ318821
VERSION AJ318821.1 GI:22208418
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE 1
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 639)
AUTHORS Herrera, L.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid: 28220, SPAIN
FEATURES
source
1. .639
Location/Qualifiers
/organism="Mycobacterium tuberculosis"
/isolate="1071-98"
/db_xref="taxon:1773"
gene 1. 639
/gene="rpoB"
CDS <1..>639
/codon_start=1
/trans_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87038.1"
/db_xref="GI:22208419"
TRANSLATION="RRTDEALLDIYKLRPGPEPTKESAOTLLENLFFKKRYDLARV
GRYKNKLGHVGEPTTSITLDEEVVATYELVRLHGGOTTMTVPGVVPEVTD
IDHFNRLRTVGLIQNIQIRGMSMERVVRMTQDVEAITPOTLINIRPAVAL
KEFFTSOLSQFMVONPLSLGTLKRRLSALGPGLSRERAGLEVRDVHPSHS"
BASE COUNT 126 a 202 c 212 g 99 t
ORIGIN
Query Match 76.9%; Score 160; DB 1; Length 639;
Best Local Similarity 85.6%; Pred. No. 1e-17;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAACGCTACGACCTGCGCGGTGCGCGTCTGCGTCTATAGGTCAACAAGAGCTCG 60
Db 101 TCAAGGAGAACGCTACGACCTGCGCGGTGCGCGTCTGCGTCTATAGGTCAACAAGAGCTCG 160
Qy 61 GCCTGAACACCGCGTCCCGGATCAGACACACACTCTGACCGAAGAGAGCTCGTCGCCCA 120
Db 161 GGCTGCATGTGCGCGGAGCCCATACGTCGTCGACGCTGACCGAAGAAGAGCTCGTGCCCA 220
Qy 121 CCATCGAGTACCTGCTCGCGCTGCACGAGGCGCACACACGATGACCGTCCCGGCGGAG 180
Db 221 CCATCGAATATCTGCTCGCGTTCGACGAGGTCAGACCGATGACCGTTCCTCCGGCGGCG 280
Qy 181 TCGAGGTGCGCGGTGGAACCGAGACAT 208
Db 281 TCGAGGTGCGCGGTGGAACCGAGACAT 308
RESULT 9
LOCUS 150706 970 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 1 from patent US 5643723.
ACCESSION 150706
VERSION 150706.1 GI:2472409
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 970)
AUTHORS Persing, D.H., Hunt, J.J., Young, K.K.Y., Felmler, T.A., Roberts, G.D.
and Whelan, A. Christian.
TITLE Detection of a genetic locus encoding resistance to rifampin in
mycobacterial cultures and in clinical specimens
JOURNAL Patent: US 5643723-A 1 01-JUL-1997;
FEATURES Location/Qualifiers
source 1. 970
/organism="unknown"
BASE COUNT 182 a 302 c 330 g 156 t
ORIGIN
Query Match 76.9%; Score 160; DB 6; Length 970;
Best Local Similarity 85.6%; Pred. No. 9.7e-18;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAACGCTACGACCTGCGCGGTGCGCGTCTGCGTCTATAGGTCAACAAGAGCTCG 60
Db 26 TCAAGGAGAACGCTACGACCTGCGCGGTGCGCGTCTGCGTCTATAGGTCAACAAGAGCTCG 85
Qy 61 GCCTGAACACCGCGTCCCGGATCAGACACACACTCTGACCGAAGAGAGCTCGTCGCCCA 120
Db 86 GGCTGCATGTGCGCGGAGCCCATACGTCGTCGACGCTGACCGAAGAGAGCTCGTGCCCA 145
Qy 121 CCATCGAGTACCTGCTCGCGCTGCACGAGGCGCCACACACGATGACCGTCCCGGCGGAG 180
Db 146 CCATCGAATATCTGCTCGCGTTGACGAGGTCAGACCGATGACCGTTCCTCCGGCGGCG 205
Qy 181 TCGAGGTGCGCGGTGGAACCGAGACAT 208
Db 206 TCGAGGTGCGCGGTGGAACCGAGACAT 233
RESULT 10
LOCUS AX111339 3534 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 2072 from Patent WO0123604.
ACCESSION AX111339
VERSION AX111339.1 GI:13927631
KEYWORDS
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE 1 (bases 1 to 3534)
AUTHORS Bergeron, M.G., Boisset, M., Huilet, A., m Nard, C., Ouellette, M.,
Picard, F.J. and Roy, P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 2072 05-APR-2001;
FEATURES Infectio Diagnostic (I.D.I.) INC. (CA)
Location/Qualifiers

|||||CCATCGAGTACCTGCTCGCGCTGCACGAGGCGCACACGATGACCGTCCCGGCGGAG 180
Qy 121
|||||CCATCGAATATCTGCTCGCGTTCGACGAGGTCAGACCGATGACCGTTCCTCCGGCGGCG 280
Db 221
|||||TCGAGGTGCGCGGTGGAACCGAGACAT 208
Qy 181
|||||TCGAGGTGCGCGGTGGAACCGAGACAT 308
Db 281
RESULT 9
LOCUS 150706 970 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 1 from patent US 5643723.
ACCESSION 150706
VERSION 150706.1 GI:2472409
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 970)
AUTHORS Persing, D.H., Hunt, J.J., Young, K.K.Y., Felmler, T.A., Roberts, G.D.
and Whelan, A. Christian.
TITLE Detection of a genetic locus encoding resistance to rifampin in
mycobacterial cultures and in clinical specimens
JOURNAL Patent: US 5643723-A 1 01-JUL-1997;
FEATURES Location/Qualifiers
source 1. 970
/organism="unknown"
BASE COUNT 182 a 302 c 330 g 156 t
ORIGIN
Query Match 76.9%; Score 160; DB 6; Length 970;
Best Local Similarity 85.6%; Pred. No. 9.7e-18;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAACGCTACGACCTGCGCGGTGCGCGTCTGCGTCTATAGGTCAACAAGAGCTCG 60
Db 26 TCAAGGAGAACGCTACGACCTGCGCGGTGCGCGTCTGCGTCTATAGGTCAACAAGAGCTCG 85
Qy 61 GCCTGAACACCGCGTCCCGGATCAGACACACACTCTGACCGAAGAGAGCTCGTCGCCCA 120
Db 86 GGCTGCATGTGCGCGGAGCCCATACGTCGTCGACGCTGACCGAAGAGAGCTCGTGCCCA 145
Qy 121 CCATCGAGTACCTGCTCGCGCTGCACGAGGCGCCACACACGATGACCGTCCCGGCGGAG 180
Db 146 CCATCGAATATCTGCTCGCGTTGACGAGGTCAGACCGATGACCGTTCCTCCGGCGGCG 205
Qy 181 TCGAGGTGCGCGGTGGAACCGAGACAT 208
Db 206 TCGAGGTGCGCGGTGGAACCGAGACAT 233
RESULT 10
LOCUS AX111339 3534 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 2072 from Patent WO0123604.
ACCESSION AX111339
VERSION AX111339.1 GI:13927631
KEYWORDS
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE 1 (bases 1 to 3534)
AUTHORS Bergeron, M.G., Boisset, M., Huilet, A., m Nard, C., Ouellette, M.,
Picard, F.J. and Roy, P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 2072 05-APR-2001;
FEATURES Infectio Diagnostic (I.D.I.) INC. (CA)
Location/Qualifiers

source	1. .3534	/organism="Mycobacterium tuberculosis"	
		/strain="Rv"	
BASE COUNT	679 a	1081 c	1188 g 586 t
ORIGIN			
Query Match	76.9%;	Score 160;	DB 6; Length 3534;
Best Local Similarity	85.6%;	Pred. No. 7.7e-18;	
Matches 178;	Conservative	0;	Mismatches 30; Indels 0; Gaps 0;
QY	1	TCAAGGAGAGCGGTACGACCTCGCGGGTGGGCGGCTACAAAGGTCAACAAAGAGCTCG	60
Db	902	TCAAGGAGAGCGGTACGACCTCGCGGGTGGGCGGCTACAAAGGTCAACAAAGAGCTCG	961
QY	61	GCCTGAACACCGCTCCCGGATCAGGACCACTCTGACCGAAGAGAGCTCGTCGCCCA	120
Db	962	GCCTGCATCTCGGAGCGCCATCGCTGTCGACGCTGACCGAAGAGAGCTCGTCGCCCA	1021
QY	121	CCATCGAGTACCTGTCGCCCTGCACGAGGCCACACCATGACCGTCCCGGGCGGAG	180
Db	1022	CCATCGAATATCTGTCGCTGCACGAGGTCAGACGATGACCGTCCCGGGCGGCG	1081
QY	181	TCGAGGTCCGGTGAACACGACGACAT	208
Db	1082	TCGAGGTCCGGTGAACACGACGACAT	1109
RESULT 11			
LOCUS	MTU12205	3853 bp	DNA linear BCT 02-MAR-2000
DEFINITION	Mycobacterium tuberculosis H37Rv RNA-polymerase beta subunit (rpoB) gene, partial cds.		
ACCESSION	U12205		
VERSION	U12205.1	GI:515684	
KEYWORDS	Mycobacterium tuberculosis.		
SOURCE	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Corynebacteriales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.		
REFERENCE	1 (bases 1 to 3853)		
AUTHORS	Imboden, P., Troller, R., Marchesi, F., Telenti, A., Bodmer, T., Cole, S., Schopfer, K. and Burkart, T.		
TITLE	The rpoB gene of Mycobacterium tuberculosis		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3853)		
AUTHORS	Imboden, P.		
TITLE	Direct Submission		
JOURNAL	Submitted (11-JUL-1994) Paul Imboden, Institute for Medical Microbiology, University of Berne, Friedbuehlstrasse 51, Berne, 3010, Switzerland		
FEATURES	Location/Qualifiers		
source	1. .3853	/organism="Mycobacterium tuberculosis"	
		/strain="H37Rv"	
gene	576. .>3853	/db_xref="taxon:1773"	
CDS	576. .>3853	/gene="rpoB"	
		/gene="rpoB"	
		/codon_start=1	
		/translation="MEEGCLADSRQSKTAASPSRQSSNNNSVPGAPNRYSAFL	
		/product="RNA-polymerase beta subunit"	
		/protein_id="AAA20242.2"	
		/db_xref="GI:7144499"	
		/translation="MEEGCLADSRQSKTAASPSRQSSNNNSVPGAPNRYSAFL	
		/product="RNA-polymerase beta subunit"	
		/protein_id="AAA21416.1"	
		/db_xref="GI:468334"	
		/translation="MLEGCLADSRQSKTAASPSRQSSNNNSVPGAPNRYSAFL	

REPLEVPGLLDVOTDSFEMWLGSPMRRESAAERGDVNPVGGLEEVLYELSPIDFSGS
MSLSFSDPRFDDVKAPVDECKDKMTVAAPLEVTAFEPINNNTGEISQTFVMDGDFPMK
TEKTFIINGTERVVSQVRSQVYFDEIDKSTDKTLHSVKYIPSRGAWLEFVDVK
RDTVGVIRDRKRQPVTVLLKALGTMSEQIVRFGSEIMRSTLEKDNVTGTDALLD
EEDVATYIELVRLHEGQTTMTVPVGGVEVETDDIDHFGNRRRLTVGELIONQIRVG
MSRNVVRERMTTQDVEAITPOTLINIRPVAAIKFEFGTSOLFQMDONPLSGLT
HKRRLSALPGGLSRERAGLEVDPVSHYGMCPLETPEGPNIGLISLVYARNP
FGFTETPVKVDGVDSDEIVLITADEDRHVVAQNSPIDAGRFRVPRVLVRRKAG
EVEVPSEVDYMDVSRQVSVATAMIPLEHDDANRALMGANMQQVAPVLRSEAP
LVGCGMELRAIDAATSSQESGVIEVSADYITVMDHNGTRRTYRMRKFAHSNHTC
ANOCPIVDAGRVEAGVVIADGCTDDGEMALGNLLVAIMPWEHNYEDAIILSNRL
VEEDVLTSHIEHEIDARIDKLGAEITRDIPNIIDEVLADDERGIVRIGAEVRD
DILVGVKTPKGETELTPEERLLRAIFCEKAREVDTSLKVPHGESCIVIGRVSRE
EDELPGVNELVVVYVAQKISDGOKLAGRHGKGVIGKILPVEDMPFLADGTPVDI
ILNTHGVRPRMNIQOILETHLGMCAHSGMKVDAAGVDPWAARLPDELLPAOPNAIV
TPVFDGAQAEALGGLSCTLPNRDGVLDVADKAMLFDRGSGEPFPYVTVGYMTIM
KLHLVDVKIHKARSTGYSMITQOPLGKAQFGGRFGEWECWAMQAGAAVTLQELL
TIKSDDTVGRKVVYEAIVKGENIPEPGIPESFKVLLKELQSLCNVLEVLSSDGAATE
REGDEDELERAAANLGINLSRNESASFEDLA"

gene 4641. .5084
CDS /gene="rpoC"
4641. .>5084
/gene="rpoC"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta'-subunit"
/protein_id="AA21417.1"
/db_xref="GI:537608"
/translation="MLDVNFFDELRIGLATAEDIRQWYGEVKKKPTINRYTLKPEKD
GLFCEKIFGTRDMWECYCGKRVFRFGIICERCQGEVTRAKVRMRMGHIELAAPVT
HIWFKGVPSRLGDLAPKDLKIIYFAAYVITSVDEMRHNEL"

BASE COUNT 969 a 1534 c 1691 g 890 t

Query Match 76.9%; Score 160; DB 1; Length 5084;
Best Local Similarity 85.6%; Pred. No. 7.2e-18;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 1 TCAAGGAGACGCTACGACCTCGCGGGTGGCCGCTCAAGGTCACAAAGACGTCG 60
Db 1966 TCAAGGAGACGCTACGACCTCGCGGGTGGCCGCTCAAGGTCACAAAGACGTCG 2025

Qy 61 GCCTGAACACGCGTCCCGCATCAGCAGCACCTCTGACCGAAGAGACGTCGTCGCCA 120
Db 2026 GCCTGCATCTGCGGAGCCATCAGCTCGTGCAGCTGACCGAAGAGACGTCGTCGCCA 2085

Qy 121 CCATCGAGTACCTGCTGCGCTGACGAGGGCCACACCATGATGACCGTCCCGGGCGGAG 180
Db 2086 CCATCGAATATCTGGTCCGCTGACGAGGGTCAGACCATGATGACCGTCCCGGGCGG 2145

Qy 181 TCGAGGTGCCGGTGAACACGACGACAT 208
Db 2146 TCGAGGTGCCGGTGAACACGACGACAT 2173

RESULT 13
AE006964
LOCUS
DEFINITION Mycobacterium tuberculosis CDC1551, linear BCT 27-APR-2001 complete genome.
ACCESSION AE006964
VERSION AE006964.1
KEYWORDS GI:13880217
SOURCE Mycobacterium tuberculosis CDC1551.
ORGANISM Mycobacterium tuberculosis CDC1551.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 19352)
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M.,

Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains
Unpublished
2 (bases 1 to 19352)
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
Direct Submission
Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
1. .19352
/organism="Mycobacterium tuberculosis CDC1551"
/strain="CDC1551"
/db_xref="taxon:83331"
/note="clinical strain"
163. .3699
/gene="MT0695"
163. .3699
/gene="MT0695"
/note="similar to GB:L27989 GB:L05910 GB:U12205 SP:P47766
PID:L49992; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA-directed RNA polymerase, beta subunit"
/protein_id="AAK44922.1"
/db_xref="GI:13880218"
/translation="MLEGILADSRQSKTAASPPSRQSSNNKSPCANRVSFAKL
REPLEVGLLDVOTDSFEMWLGSPMRRESAAERGDVNPVGGLEEVLYELSPIDFSGS
MSLSFSDPRFDDVKAPVDECKDKMTVAAPLEVTAFEPINNNTGEISQTFVMDGDFPMK
TEKTFIINGTERVVSQVRSQVYFDEIDKSTDKTLHSVKYIPSRGAWLEFVDVK
RDTVGVIRDRKRQPVTVLLKALGTMSEQIVRFGSEIMRSTLEKDNVTGTDALLD
EEDVATYIELVRLHEGQTTMTVPVGGVEVETDDIDHFGNRRRLTVGELIONQIRVG
MSRNVVRERMTTQDVEAITPOTLINIRPVAAIKFEFGTSOLFQMDONPLSGLT
HKRRLSALPGGLSRERAGLEVDPVSHYGMCPLETPEGPNIGLISLVYARNP
FGFTETPVKVDGVDSDEIVLITADEDRHVVAQNSPIDAGRFRVPRVLVRRKAG
EVEVPSEVDYMDVSRQVSVATAMIPLEHDDANRALMGANMQQVAPVLRSEAP
LVGCGMELRAIDAAGVVIADGCTDDGEMALGNLLVAIMPWEHNYEDAIILSNR
CANOCPIVDAGRVEAGVVIADGCTDDGEMALGNLLVAIMPWEHNYEDAIILSNR
VEEDVLTSHIEHEIDARIDKLGAEITRDIPNIIDEVLADDERGIVRIGAEVRD
DILVGVKTPKGETELTPEERLLRAIFCEKAREVDTSLKVPHGESCIVIGRVSRE
EDELPGVNELVVVYVAQKISDGOKLAGRHGKGVIGKILPVEDMPFLADGTPVD
ILNTHGVRPRMNIQOILETHLGMCAHSGMKVDAAGVDPWAARLPDELLPAOPNAIV
TPVFDGAQAEALGGLSCTLPNRDGVLDVADKAMLFDRGSGEPFPYVTVGYMTI
MKHLVDVKIHKARSTGYSMITQOPLGKAQFGGRFGEWECWAMQAGAAVTLQELL
TIKSDDTVGRKVVYEAIVKGENIPEPGIPESFKVLLKELQSLCNVLEVLSSDGAATE
REGDEDELERAAANLGINLSRNESASFEDLA"

gene 3744. .7694
CDS /gene="MT0696"
3744. .7694
/gene="MT0696"
/note="similar to SP:P37871; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA-directed RNA polymerase, beta-prime subunit"
/protein_id="AAK44922.1"
/db_xref="GI:13880219"
/translation="MLDVNFFDELRIGLATAEDIRQWYGEVKKKPTINRYTLKPEKD
GLFCEKIFGTRDMWECYCGKRVFRFGIICERCQGEVTRAKVRMRMGHIELAAPVT
HIWFKGVPSRLGDLAPKDLKIIYFAAYVITSVDEMRHNEL"
AVEDQRDGELEPAAKLEADLAELAEAGAKADARRKVRDGEREMQIRRAQREDLR
LEDIMSTFTKLAPQLIVDENLRELVDRTGETFTGANGAESOKLLENFDIDAEAS
LRDVRNGKQKRLRUKRVAAVQSGNSPMGMVLDVAVPVPIDELRPMVQDGGGR
FATSDLDLRYRNNRNLKRLIDGAPETIYNNEKRMILQESVDLFDNGRGRPVT
GPNRPLKSLDLKQKQGRFRONLLKRVDSGRVIVYGPOLKHOCGLPKLMALE
LFKPFVKRLVDLHNAQNIKSAKRMVRORPQVMDVLEEVIAEHPVLLNAPRTLRL
IQAEFPMVESKATQLHPLVCEAFNADFDGDMVHLPLSAEAQAERAILMLSSNLT

gene
CDS
/transl_table=11
/product="acyl-CoA dehydrogenase, putative"
/protein_id="AAK44926.1"
/db_xref="GI:13880223"
/translation="MSDTHVTVNVPPLNYPASSPVLIIEALIOEGOMGLDFVNEV
GAISCAOQWGLADNRPLIHTHDAYGYRDEVEDPAYHELMRTAITHGHAAP
WADPRGHVRAAKTSVTPVEPHICISMTYAVVPALRINSELAAYEPPLTSREY
DELPAPHTTAGITAGMSWTEKQSGDVRAGTQATPNADGSLTGHKMFPSAMCD
IFVLQAPADKLSCLFLPRVLPDSTNRMFQRLKDLKLNHANASSEVEYDGAVALV
GEGRGVPTIIEVNLTKLDCALASATSMRTGLTRAVHQAHRKAFGAYLIDPLMRN
VLADLAVEAEATIVAMRAGATNAVGRNETEALLRRIGLAAKAYVCKRSTAAAE
ALECLGGNGVEDSCMPRLYREAPLWGTWEGSGNVSALDTLRANATPACVEVLFDEL
ARSAGNCR2LDGSHVERLPPGLDLDITIGYRARKTAEDICLALQSLVLRHGHGPAVEA
FLATKLLCGWGGANGTTPAGLDLAPILERALVKG"
13498..14436
/gene="MT0702"
13498..14436
/gene="MT0702"
/note="similar to GP:3885480; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="enoyl-CoA hydratase/isomerase family protein"
/protein_id="AAK44927.1"
/db_xref="GI:13880224"
/translation="MTHAIRVDFDNLTWYTVTGRIARITFNRPEKGNALIAIDTPL
ELSLVERADLDPGVHVILVSGRGEFCAGFDLSAYAESSSTGGGAYOGTVLDGKT
QAVNHLPNQPMIDYOMSRFVRGASLMHADKPTVYKTHGYCVAAGTDLIALHDQ
VTAADAKIGWPTVRVGVPAAGLWHRLLQDQAKRLFTGDCITGAQAAMGLAVEA
PEPADLDERLVARIAALPVNOLIMVKLALNSALLQOQVATSRMSTVFDGAARHT
PEGHAFVADAVEHGRDVARRRDEPFQDYGQASRV"
14439..15161
/gene="MT0703"
14439..15161
/gene="MT0703"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"

gene
CDS
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44923.1"
/db_xref="GI:13880220"
/translation="MFDSSAAITNPGHANASAMERSGLLECVCAGLDQPFGEFTADKL
NPDRGSSRRVPRROADGGIATHVHGGGQSGQAGVVPQRMHGFEPALAMQDRLIH
GEQTNRIQAQFRFCVCSPT"
complement(8058..9972)
/gene="MT0698"
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact; identified by
Glimmer2; putative; conserved hypothetical protein,
authentic frameshift"
10167..10925
/gene="MT0699"
10167..10925
/gene="MT0699"
/note="identified by match to PFAM protein family HMM
PF01261"
/codon_start=1
/transl_table=11
/product="AP endonuclease, family 2"
/protein_id="AAK44924.1"
/db_xref="GI:13880221"
/translation="MLIGSHVSPTDPLAAAEAGDVVOIFLGNPQSWKAPKPRDAA
ALKAATLPIYHAPYLINLASANNRVRIPSRLIQETCAAAADIGAAAIVHGGHVD
DNDIDRFGWRKALDLRETEVPVYLENTAGDHAMARRFDIARLWDVIGDTGIFGC
LDTCHTWAAGEALTDAVDRIKAITGRIDLHVCNDSRDEAGSGRDRHANLGSQIDPDL
LVAAVRAAGAPVICETADQGRKNDINFLERTGS"
10957..11799
/gene="MT0700"
10957..11799
/gene="MT0700"
/note="similar to GB:U00012 PID:466863; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hydrolase/esterase, putative"
/protein_id="AAK44925.1"
/db_xref="GI:13880222"
/translation="MLRRVAILLAVALFAGCGGTRLAAGFGNGSNVHTLDVDCAGR
SVRLXKPLVGLSPALVLMHGGEGSKAERSYGNDELASEKFLVAYPDGYHRAWN
ANGGCGGPAREGVDDIGFVRVAVDIANNYSIDPARVYVTGMSNGAIMSYTLACNT
SIFAAIGVSGTQDPCQSPRPVSVIHITADPLRVHGGPGAFADIGPPVVDLN
AFWRVLRGVTQDITTEGPVTTSGATCADNRRVLLTVDADGHRWPSFATQTLWRFFA
AHFR"
11859..13487
/gene="MT0701"
11859..13487
/gene="MT0701"
/note="similar to SP:P33224 GB:L20915 PID:457172
PID:457174 PID:537028; identified by sequence similarity;
putative"
/codon_start=1

Query Match 76.9%; Score 160; DB 1; Length 19352;
Best Local Similarity 85.6%; Pred. No. 5.6e-18;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1 TCAAGGAGAAGCGCTACGACCTCGCGCGGTGGCGGTCTACAGGTCAACAAGAGCTCG 60
DB 1064 TCAAGGAGAAGCGCTACGACCTCGCGCGGTGGCGGTCTACAGGTCAACAAGAGCTCG 1123
QY 61 GCCTGAACACCGCGTCCCGGATCAGCAGGACCCACTCTGACCGAAGAGCAGCTCGTGGCCA 120
DB 1124 GCCTGATGTCGCGGAGCCATCAGCTCGTCCGACCTGACCGAAGAGCAGCTCGTGGCCA 1183
QY 121 CCATCGAGTACCTGGTCCGCTGACGAGGCGCCACACCATGATGACCTCCGCGGCGAG 180
DB 1184 CCATCGAATATCTGTCGCTTGCACGAGGTTCAGACCATGATGACCTCCGCGGCGCG 1243
QY 181 TCGAGGTGCGGTGGAACCGAGCAGCAT 208
DB 1244 TCGAGGTGCGGTGGAACCGAGCAGCAT 1271

RESULT 14
MTCI376
LOCUS MTCI376 19770 bp DNA linear BCT 03-AUG-2001
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 32/162.
ACCESSION Z95972 ALI23456
VERSION Z95972.1 GI:3261790
KEYWORDS
SOURCE
ORGANISM
Mycobacterium tuberculosis H37Rv.
Mycobacterium tuberculosis H37Rv
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE
AUTHORS

1 (bases 1 to 19770)
Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,
Harris, D., Gordon, S.V., Eiglmeyer, K., Gas, S., Barry III, C.E.,
Tekala, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,
Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S.,
Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J.,
Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A.,
Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skellon, S.,
Squares, S., Squires, R., Sulston, J.E., Taylor, K., Whitehead, S. and
Barrell, B.G.
Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence
Nature 393 (6685), 537-544 (1998)
98295987
PUBMED
9634230
2 (bases 1 to 19770)
Parkhill, J.
Direct Submission
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
tuberculosis sequencing and mapping teams: Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 27, 1998 this sequence version replaced gi:2143285.

COMMENT

Notes:
Details of M. tuberculosis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have
been renumbered from the original cosmid submissions but the old
gene designations are in brackets after the new gene numbers.
Gene prediction was based on a Hidden Markov Model of TB genes
implemented in TBparse (Krogh) supplemented with visual inspection
of positional base preference in codons, especially where there is
an increase in the observed/expected third position G + C.
CAUTION: In some cases we may not have predicted the correct
initiation codon. Where possible we choose an initiation codon
(atg, gtg, or ttg) which is preceded by an upstream ribosome
binding site sequence (optimally 5-13bp before the initiation
codon). If this cannot be identified we choose the most upstream
initiation codon.

FEATURES
source

Location/Qualifiers
1. .19770
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
/clone="1376"
57. .61
/note="agga, possible rbs upstream of Rv0654"
68. .1573
/gene="Rv0654"
68. .1573
/gene="Rv0654"
/note="Rv0654", (MTCI376.22), len: 501. unknown, FASTA
score: Q53353 LIGNOSTILBENE-ALPHA, BETA-DIOXYGENASE (485
aa) opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity
in 523 aa overlap). Also similar to M. tuberculosis protein
MTCY21C12.07c (29.5% identity in 522 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0654"
/protein_id="CAB09380.1"
/db_xref="GI:2143307"
/db_xref="SPTREMBL:O06785"
/translation="MTAAQAESQNPYLEGLAPVSTVATDLPVTGRIPEHLDRGY
LRGNPNPAVEVPATYHFTGDAMVGVALRDGKARWRNKRVRPAPCAALGPISA
RHPRTGIIEGPNLVLTAGRLVLAEGVYNTLELDELDTVGPCDFDGLHGYYT
APRPDPHGLHAVSHVSPARGHVOYSVIGTDGHARTVDIEVAGSPMHFSJLTDN
YVYVLDLPFTFPMQVPASVPRLQPARLVISGLVRLNIPDPFAALGNRMQGHSD
RLPYAWNPSPYPARGVMPREGNEDVRWFDEIPCYVYHPLRVNAYSECRNGAEVLVDVV
RYSRMFDRRGPGGDRPSLDRTWTINLATGATVACEDRDRAQEPFRINETLVGPHR
FATYVIGEGELVGAGAAALSTPLYKQDCVCTGSSVATLDPOLLIGEMVFVNPSPRAE
DGGILMGYWHHGRGEGQLLLDQAOTLESATVHLHPQRPVPMGFHGNWAPT"
1585. .2664

gene

CDS

/gene="Rv0655"
1585. .2664
/gene="Rv0655"
/note="Rv0655, (MTCI376.21), len: 359, abc transporter,
FASTA score: YRBF_ECOLI P45393 hypothetical abc
transporter atp-binding (269 aa) opt: 644 z-score: 721.8
E(): 3.4e-33 (38.5% identity in 244 aa overlap): contains
PS00017 ATP/GTP-binding site motif A, PS00211 ABC
transporters family signature, highly similar to M. leprae
MKL_MYCLE P30769 possible ribonucleotide transport atp-
(347 aa) opt: 2021 z-score:2244.4 E(): 0, (92.2% identity
in 335 aa overlap). Also similar to many other M.
tuberculosis ABC transporters eg. MTCY253.24 (33.6%
identity in 241 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0655"
/protein_id="CAB09379.1"
/db_xref="GI:2143306"
/db_xref="SWISS-PROT:O06784"
/translation="MRYSDSYHTTGRWQPRASTGFPMPGVSIEVNGLTKSFSSRIWE
DVTLTIPAGEVSVLLGPGSGTKSVFLKSLIGLLRPERGSIIDGTDIIECSAKELYEI
RTLFGVLQDQKALFGSMNLNDYNTAPPLREHTPKKKESEIRDIYMEKALVGLGGDKKF
PGEISGGHKKRAGLARALVLDPLIICLDEPDSGLDPVATYLSQLIMDINAOIATIL
LYVINIARTVPDNNMGLFRKHLVFMFGPREVLTSDPEVYVQLFNGRRIRGIPGNSEE
KDEATMAEEQALLDAGHHAGGVEETGEPVQISATPGMPERKAVARQARVREMLHTL
PKKAAAILDLLEGTHKYAVHEIGQ"
1762. .1785
/gene="Rv0655"
/note="PS00017 ATP/GTP-binding site motif A"
2074. .2118
/gene="Rv0655"
/note="PS00211 ABC transporters family signature"
complement(3052. .3435)
/gene="Rv0656c"
complement(3052. .3435)
/gene="Rv0656c"
/note="Rv0656c, (MTCI376.20, unknown), len: 127 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0656c"
/protein_id="CAB09378.1"
/db_xref="GI:2143305"
/db_xref="SPTREMBL:O06783"
/translation="MAAATTTGTHRGLELRAAQRAVGSCEPQRAEFCRSARNADEFDQ
MSRMFGDVPDPVPKSVWRWIDSAQHLRAGAVGALSVVDLLICDTAAAGLVLH
DDADYELAEARHLPDIRVRVVSADD"
complement(3530. .3685)
/gene="Rv0657c"
complement(3530. .3685)
/note="Rv0657c, (MTCI376.19), unknown, len: 51 aa; similar
to several other M. tuberculosis hypothetical proteins eg.
YW08_MYCTU Q10848 hypothetical 8.9 kd protein cy39.08c (80
aa) fasta scores: opt: 107 z-score: 182.3 E(): 0.0038,
45.8% identity in 48 aa overlap. Also similar to MTCY48_5
and AL020958|SC4H8_7 Streptomyces coelicolor cosmid 4H8 (66
aa), 41.0% identity in 39 aa overlap."
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0657c"
/protein_id="CAB09377.1"
/db_xref="GI:2143304"
/db_xref="SPTREMBL:O06782"
/translation="MSVTQIDLDELADVNRIAAVHTKKKAVNLAMRDYVERFRRIE
ALARSRE"
complement(3691. .3696)
/note="possible RBS upstream of Rv0657c"
complement(3761. .4477)
/gene="Rv0658c"
complement(3761. .4477)
/gene="Rv0658c"
/note="Rv0658c, (MTCI376.18), len: 238, unknown,

Db 1248 TCGAGTTCGCCGTCGAGTCGCACGACAT 1275

Search completed: November 13, 2002, 01:26:47
Job time : 660.723 secs

GenCore version 5.1.3
Copyright (c) 1993 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 ; Search time 1189.12 Seconds
(without alignments)
3037.202 Million cell updates/sec

Title: US-09-697-123b-13

Perfect score: 223

Sequence: 1 tcaaggagaagcgctacgac.....ccggtggaaccgacgacat 223

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pin:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	123.8	55.5	1282	9	AI770311
2	42.4	19.0	431	13	BI097733
3	42.4	19.0	507	13	BI096671
C 4	42.4	19.0	531	13	BM348454
C 5	42.4	19.0	579	13	BM340507
6	42.4	19.0	586	12	BG842655

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C 7	42.4	19.0	593	13	BM079854
C 8	42.4	19.0	597	13	BM333485
C 9	42.4	19.0	606	13	BM337568
C 10	42.4	19.0	610	13	BM073467
C 11	42.4	19.0	613	13	BM340956
C 12	42.4	19.0	622	13	BM337575
C 13	42.4	19.0	631	13	BM332944
C 14	42.4	19.0	659	13	BM266851
C 15	42.4	19.0	659	13	BM337566
C 16	42.4	19.0	687	13	BM349198
C 17	42.4	19.0	701	13	BM075264
C 18	41.6	18.7	530	13	BM336791
C 19	40.4	18.1	511	10	AV628007
C 20	40.4	18.1	543	10	AV631761
C 21	40.4	18.1	554	10	AV631497
C 22	40.4	18.1	725	14	BM023567
C 23	40.2	18.0	524	13	BM336777
C 24	40	17.9	546	12	BF257512
C 25	39.4	17.7	494	13	BJ318566
C 26	39.4	17.7	649	13	BJ312832
C 27	39.2	17.6	433	10	AM923980
C 28	39.2	17.6	531	12	BG841942
C 29	39.2	17.6	536	10	AM679799
C 30	39.2	17.6	546	10	AV433406
C 31	39.2	17.6	604	13	BI721593
C 32	39.2	17.6	655	9	AU170926
C 33	39	17.5	300	9	AU176504
C 34	39	17.5	363	9	AU162829
C 35	38.8	17.4	632	13	BM266831
C 36	38.4	17.2	562	14	BQ752716
C 37	38	17.0	459	9	AU033206
C 38	38	17.0	461	14	D49085
C 39	37.8	17.0	533	10	BE405789
C 40	37.8	17.0	566	13	BM076409
C 41	37.8	17.0	594	13	BJ317721
C 42	37.8	17.0	617	13	BJ323305
C 43	37.8	17.0	710	12	BG321153
C 44	37.8	17.0	1057	12	BF345788
C 45	37.8	17.0	1856	11	AY109418

ALIGNMENTS

RESULT 1	AI770311/c	AI770311	1282 bp	linear	EST 24-JAN-2000
LOCUS	42 Mycobacterium anaerobic stationary phase library Mycobacterium smegmatis CDNA, mRNA sequence.	AI770311			
DEFINITION	AI770311				
ACCESSION	AI770311.1	GI:6742680			
VERSION	EST.				
KEYWORDS	Mycobacterium smegmatis.				
SOURCE	Mycobacterium smegmatis				
ORGANISM	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.				
REFERENCE	1 (bases 1 to 1282)				
AUTHORS	Murugasu-Oei, B., Tay, A. and Pick, T.				
TITLE	Upregulation of stress response genes and ABC transporters in anaerobic stationary-phase Mycobacterium smegmatis				
JOURNAL	Mol. Gen. Genet. 262 (4-5), 677-682 (1999)				
MEDLINE	20092472				
COMMENT	Contact: Murugasu-Oei, B. Mycobacterium Laboratory Institute of Molecular and Cell Biology 30 Medical Drive, Singapore 117609, Republic of Singapore Tel: 65 874 3011 Fax: 65 779 1117 Email: mcbom@imcb.nus.edu.sg Insert Length: 1282 Std Error: 0.00 Seq primer: T3 Forward; T7 Backward. Location/Qualifiers 1. .1282				

/organism="Mycobacterium smegmatis"
 /strain="MC2155"
 /db_xref="taxon:1772"
 /clone_lib="Mycobacterium anaerobic stationary phase library"
 /lab_host="E. coli XL1-Blue MRF"
 /note="Vector: Lambda ZAP II; Bacilli were disrupted using a Mini bead beater (Biospec). RNA was isolated using the RNeasy protocol (Qiagen). Purified RNA was subjected to 2 rounds of digestion with RNase-free DNase I (Promega). DNase I was heat-inactivated at 75°C for 5 min, and removed by using RNeasy columns followed by phenol extraction and ethanol precipitation. The RNA preparations were confirmed to be free of genomic DNA contamination by carrying out PCR and RT-PCR using the Access kit (Promega) and primers specific for the histone-like protein gene *hlp* (Lee et al., 1998). cDNA was synthesized using random hexamer primers (Promega) and Stratagene's cDNA synthesis kit. cDNA fragments were ligated into lambda ZAP II vector and packaged in vitro using Stratagene reagents."
 BASE COUNT 230 a 395 c 405 g 252 t
 ORIGIN

Query Match 55.5%; Score 123.8; DB 9; Length 1282;
 Best Local Similarity 76.7%; Pred. No. 1.7e-19;
 Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
 QY 1 TCAAGAGAGCGCTACGACCTGGCGCGCTACAGGTCACCAAGAAGCTGG 60
 Db 236 TCAAGAGAGCGCTACGACCTGGCGCGCTACAGGTCACCAAGAAGCTGG 177
 QY 61 GCCTGACACCATCATCGGATCACACACGACGCTGACCGAAGAAGCTGCGCCA 120
 Db 176 GCCTGAACCGGCGAAGCGGATCACAGCTCGACGCTGACCGAAGAGGCTGCGCGGA 117
 QY 121 CCATCGAGTACCTGGTTCGCTCCCTCCACACGCTCTCAGGTTGGCCAGCCCGCTTANGA 180
 Db 116 CCATCGAGTACCTGGTTCGCTCCACAGG-----GTGAGACTTCGATGA 72
 QY 181 CTGTCCCGCGGGGTGCGAGTGGCGGTGGAACCGACGACAT 223
 Db 71 CGTCCCGGGTGGCGTTCGAGGTTCCCGTCGAGTTCGACGACAT 29

RESULT 2
 BI097733
 LOCUS 949017G04.y1 949 - Juvenile leaf and shoot cDNA from Steve Moose
 DEFINITION 431 bp mRNA linear EST 26-JUN-2001
 zeae mays cDNA, mRNA sequence.
 ACCESSION BI097733
 VERSION BI097733.1 GI:14569410
 KEYWORDS EST.
 SOURCE zeae mays.
 ORGANISM zeae mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 431)
 Walbot,V.
 Maize ESTs from various cDNA libraries sequenced at Stanford University
 Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 949017 row: G column: 04.
 Location/Qualifiers
 1. .431
 FEATURES source

/organism="Zea mays"
 /cultivar="W64A"
 /db_xref="taxon:4577"
 /clone_lib="949 - Juvenile leaf and shoot cDNA from Steve Moose"
 /tissue_type="Immature leaf primordium and vegetative meristem"
 /dev_stage="4 stages from 3-13 days after imbibing"
 /lab_host="E. coli XL0LR"
 /note="Organ: Juvenile vegetative shoots; Vector: pAD-GAL4-2.1; Site_1: EcoRI; Site_2: XhoI; Equal amounts of total RNA by weight from 4 tissue sources (see below) were pooled, polyA+ RNA isolated, and cDNA synthesized for EcoRI (5') and XhoI (3') directional cloning into lambda Hybrizap vector from Stratagene. Tissue Sources: 1. Whole shoots 3 days after sowing/imbibing in wet soil. 2. Basal 1.5 cm shoots 6 days after sowing - includes yellow portions of developing leaves 1-5, primordia from 6-8, and the vegetative apex. 3. Non-green portions of developing leaves 4-5 and the vegetative apex, including adult leaf primordia, 9 days after sowing. 4. Partially expanded and greening leaves 4-5 at 13 days after sowing."
 BASE COUNT 102 a 143 c 111 g 75 t
 ORIGIN

Query Match 19.0%; Score 42.4; DB 13; Length 431;
 Best Local Similarity 58.9%; Pred. No. 2.4;
 Matches 73; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
 QY 4 AGGAGAAGCGCTAGGACCTGGCGCGCTGCGGTCACCAAGAAGCTGGGCC 63
 Db 105 AGCTGATGCCCGACGACCTGGCCACCGATGCCGGAAGGAGGTCCACCAAGGTCGGCC 164
 QY 64 TGAACACCGGATCATCGGATCACCCACGACGCTGACCGAAGAAGCGTCTGCCACCA 123
 Db 165 TCGTCTACGAGCTCACATCGCCCAAGAAGATCGCGTTGATGAGGAGATCTTCGTGG 224
 QY 124 TCGA 127
 Db 225 TCGA 228

RESULT 3
 BI096671
 LOCUS 949019B12.y1 949 - Juvenile leaf and shoot cDNA from Steve Moose
 DEFINITION 507 bp mRNA linear EST 22-JUN-2001
 zeae mays cDNA, mRNA sequence.
 ACCESSION BI096671
 VERSION BI096671.1 GI:14529538
 KEYWORDS EST.
 SOURCE zeae mays.
 ORGANISM zeae mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 507)
 Walbot,V.
 Maize ESTs from various cDNA libraries sequenced at Stanford University
 Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 949019 row: B column: 12.
 Location/Qualifiers
 1. .507
 FEATURES source

REFERENCE AUTHORS TITLE JOURNAL COMMENT	FEATURES source	BASE COUNT
<p>Clade: Panicoideae; Andropogoneae; Zea.</p> <p>1 (bases 1 to 579)</p> <p>Wen,T.J., Qiu,F., Qiu,L., Ashlock,D.A and Schnable,P.S.</p> <p>Expressed sequence Tags from B73 Maize: various stages and tissues including seedlings treated with a variety of hormones</p> <p>Unpublished (2001)</p> <p>Contact: Patrick S. Schnable</p> <p>Schnable Laboratory</p> <p>Iowa State University</p> <p>G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA</p> <p>Tel: 515-294-0975</p> <p>Fax: 515-294-2299</p> <p>Email: schnable@iastate.edu</p> <p>Individual basecall and confidence value were assigned using the Phred software,</p> <p>(http://depts.washington.edu/ventures/collabtr/direct/index.htm#B73rt). Overall sequence quality assessment and vector trimming were conducted using the Lucy software (">http://www.tigr.org/softlab/>). Lucy parameters were set to ensure an overall trimmed quality of 97.5% or better without any vector fragments in the chosen high-quality region of each sequence. Low-quality bases between the poly-T and the high-quality region were replaced with N's to serve as spacers.</p> <p>PCR Primers</p> <p>FORWARD: primer T7-1 (AA TAC GAC TCA CTA TAG)</p> <p>BACKWARD: primer T3 (ATT AAC CCT CAC TAA AG)</p> <p>Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).</p>	<p>Location/Qualifiers</p> <p>1. 579</p> <p>/organism="Zea mays"</p> <p>/cultivar="B73"</p> <p>/db_xref="taxon:4577"</p> <p>/clone="MEST323-F04"</p> <p>/clone_1lib="ISUM5-RN"</p> <p>/tissue_type="mixed"</p> <p>/lab_host="DH10B"</p> <p>/note="Vector: pT7T3PAC; Site_1: EcoRI; Site_2: NotI;</p> <p>Tissues: Germinated seed and seedlings (1, 2, 8, 11 DAG), Mixed mature tissues (17, 21, 38, 69, 77 DAG), Kernels (3, 5, 10, 15, 20, 25, 30, DAP), Adventitious roots (65 DAG), Tassel (3-39 cm, 53 and 56 DAG), Immature ear (0.2-3.0 cm, 53, 56, 59 DAG), Husk (73 DAG), Silk, unpollinated first ear, ear shank, etiolated seedlings, callus, cycloheximide-treated callus, Anaerobic treated seedlings, NAA (a-Naphthalene acetic acid)-treated seedlings, Kinetin-treated seedlings, ACPG (1-aminocyclopropane-1-carboxylic acid)-treated seedlings, Brassinolide-treated seedlings, ABA (Abscissic acid)-treated seedlings, GA (Gibberellic acid)-treated seedlings, JA (Jasmonic acid)-treated seedlings. ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo-dT primer (5' AACTGAAGAAATTCGGCGCGAGAAATTTTTTTTTTTTTT). The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with NotI and size-selected. The resulting molecules were directionally cloned into the EcoRI and NotI sites of the pT7T3PAC vector. The library then went through one round of normalization to Cor value of 5 based on the Methods of Marcelo Bento Soares (Genome Research 6: 791-806, 1996)."</p> <p>123 a 145 c 147 g 163 t 1 others</p>	

	Query Match	19.0%	Score 42.4;	DB 13;	Length 579;
	Best Local Similarity	58.9%	Pred. No. 2.5;		
	Matches	73;	Conservative	0;	Mismatches 51; Indels 0; Gaps 0;
QY	4	AGGAGAGCGCTACGACCTGGCCCGCTCGCCCGCTACAAAGTCAACAAAGCTGGGCC	63		
Db	509	AGCTGATGCGCCAGACACTGGCCACCGATCGCGCAAGAGGTCCACCAAGAGTCTCGCC	450		

QY 64 TGAACACCGATCATCCGATCATCACCCACGACGCTGACCCGAAGAAGACGTCTCTGCCACCA 123
 |||
 Db 449 TCGTCTACCAAGCTCAACATCGCCGCCAAGAAGATCGCGCTTGTGATGAGGAGATCTTCGTGG 390
 |||
 QY 124 TCGA 127
 |||
 Db 389 TCGA 386
 |||

RESULT 6
 BG842655 586 bp mRNA linear EST 29-MAY-2001
 DEFINITION MEST34-E06.T7-1 ISUM3-TL Zea mays cDNA clone MEST34-E06 5', mRNA
 sequence.
 ACCESSION BG842655
 VERSION BG842655.2 GI:14244717
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 586)
 Qiu,F., Cui,F., Guo,L., Ashlock,D.A., Wen,T.J. and Schnable,P.S.
 TITLE Expressed Sequence Tags from B73 Maize Seedlings and Silks
 JOURNAL Unpublished (2001)
 COMMENT On May 25, 2001 this sequence version replaced gi:14208977.
 Contact: Patrick S. Schnable
 Schnable Laboratory
 Iowa State University
 G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
 Tel: 515-294-0975
 Fax: 515-294-2299
 Email: schnable@iastate.edu

PCR Primers
 FORWARD: T7-1 (AA TAC GAC TCA CTA TAG)
 BACKWARD: T3 (ATT AAC CCT CAC TAA AG)
 Seq primer: primer T7-1 (AA TAC GAC TCA CTA TAG).
 Location/Qualifiers
 1..586
 /organism="Zea mays"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone="MEST34-E06"
 /clone_lib="ISUM3-TL"
 /tissue_type="Seedling and silk"
 /lab_host="DH10B"
 /note="Vector: pT73PAC; Site 1: EcoRI; Site 2: NotI; ds-cDNA molecules were generated as follows. First-strand cDNA was prepared from oligo-dT selected mRNA by priming with a NotI oligo-dT primer (5' AACTGAAGATTCGGCGCGCAGGAATTTTTTTTTTTTTTTTTT). The resulting DNA:RNA hybrid was treated with RNase H and used as a template for DNA PolI-catalyzed second strand synthesis. After the addition of EcoRI adaptors, the ds-cDNAs were digested with NotI and size-selected. The resulting molecules were directionally cloned into the EcoRI and NotI sites of the pT73PAC vector."

BASE COUNT 148 a 172 c 143 g 123 t
 ORIGIN

Query Match
 Best Local Similarity 19.0%; Score 42.4; DB 12; Length 586;
 Matches 73; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 4 AGGAGAACCGCTACGACCTGGCCGCGCTCGCGCGCTACAAGGTCACCAAGAGCTGGGCC 63
 |||
 Db 152 AGCTGATGCGCCAGCACCTGGCCACCCATCGCGCAAGGAGGTCAACCAAGAGCTGGGCC 211
 |||

QY 64 TGAACACCGATCATCCGATCATCACCCACGACGCTGACCCGAAGAAGACGTCTCTGCCACCA 123
 |||
 Db 212 TCGTCTACCAAGCTCAACATCGCCGCCAAGAAGATCGCGCTTGTGATGAGGAGATCTTCGTGG 271
 |||

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 12, 2002, 16:50:46 : Search time 636.023 Seconds
(without alignments)
9517.553 Million cell updates/sec

Title: US-09-697-123b-8
Perfect score: 208
Sequence: 1 tcaagagaagcgctacgac.....ccggtgaaaccgacgacat 208

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :		GenEmbl:*
1:	gb_ba:*	
2:	gb_htg:*	
3:	gb_in:*	
4:	gb_om:*	
5:	gb_ov:*	
6:	gb_pat:*	
7:	gb_ph:*	
8:	gb_pl:*	
9:	gb_pr:*	
10:	gb_ro:*	
11:	gb_sts:*	
12:	gb_sy:*	
13:	gb_un:*	
14:	gb_vl:*	
15:	em_ba:*	
16:	em_fun:*	
17:	em_hum:*	
18:	em_in:*	
19:	em_mu:*	
20:	em_om:*	
21:	em_or:*	
22:	em_ov:*	
23:	em_pat:*	
24:	em_ph:*	
25:	em_pl:*	
26:	em_ro:*	
27:	em_sts:*	
28:	em_un:*	
29:	em_vl:*	
30:	em_htg_hum:*	
31:	em_htg_inv:*	
32:	em_htg_other:*	
33:	em_htg_mus:*	
34:	em_htg_pln:*	
35:	em_htg_rod:*	
36:	em_htg_man:*	
37:	em_htg_vrt:*	
38:	em_sy:*	
39:	em_htgo_hum:*	
40:	em_htgo_mus:*	
41:	em_htgo_other:*	

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	166.4	80.0	610	1	MTU318818	AJ318818 Mycobacte
2	166.4	80.0	610	1	MTU318819	AJ318819 Mycobacte
3	166.4	80.0	616	1	MTU318813	AJ318813 Mycobacte
4	166.4	80.0	618	1	MTU318815	AJ318815 Mycobacte
5	166.4	80.0	618	1	MTU318817	AJ318817 Mycobacte
6	166.4	80.0	633	1	MTU318814	AJ318814 Mycobacte
7	166.4	80.0	637	1	MTU318816	AJ318816 Mycobacte
8	166.4	80.0	639	1	MTU318821	AJ318821 Mycobacte
9	166.4	80.0	970	6	150706	Sequence 1
10	166.4	80.0	3534	6	AX111339	Sequence
11	166.4	80.0	3853	1	MTU12205	U12205 Mycobacteri
12	166.4	80.0	5084	1	MSGRPOB	L27989 Mycobacteri
13	166.4	80.0	19352	1	AE006964	AE006964 Mycobacte
14	166.4	80.0	19770	1	MTC1376	295972 Mycobacteri
15	156.8	75.4	3752	1	MSU24494	U24494 Mycobacteri
16	142.4	68.5	3447	6	AR067447	AR067447 Sequence
17	142.4	68.5	37617	1	MLB1790G	Z14314 M.leprae ge
18	142.4	68.5	348950	1	MLEPRTN7	AL583923 Mycobacte
19	118.8	57.1	3941	1	AF242549	AF242549 Mycolato
20	80.4	38.7	32923	1	SCD82	AL160431 Streptomy
21	80	38.5	3495	6	AX120631	AX120631 Sequence
22	80	38.5	328050	1	AP005275	AP005275 Corynebac
23	80	38.5	349980	6	AX127144	AX127144 Sequence
24	47.2	22.7	139467	1	D90905	D90905 Synchocyst
25	46.6	22.4	204050	1	AL646070	AL646070 Ralstonia
26	45.8	22.0	204050	1	AL646073	AL646073 Ralstonia
27	45.8	22.0	347660	1	AP002994	AP002994 Mesorhizo
28	43.2	20.8	3027	3	AY061587	AY061587 Drosophil
29	43.2	20.8	3920	3	AY119444	AY119444 Drosophil
30	43.2	20.8	4470	3	AF001796	AF001796 Drosophil
31	43.2	20.8	125766	2	ACU13137	ACU13137 Drosophil
32	43.2	20.8	171376	3	AC007929	AC007929 Drosophil
33	43.2	20.8	220842	3	AE003745	AE003745 Drosophil
34	43	20.7	186159	3	AC008202	AC008202 Drosophil
35	42.6	20.5	18471	1	SCBAC17F8	AL596030 Streptomy
36	41.6	20.0	349116	1	AP003003	AP003003 Mesorhizo
37	41.4	19.9	1246	1	AF469953	AF469953 Streptomy
38	41.2	19.8	152203	2	OSJN00166	AF62967 Oryza sat
39	41	19.7	195859	14	AF281817	AF281817 Tupaia he
40	40.8	19.6	204050	1	AL646073	AL646073 Ralstonia
41	40.6	19.5	7870	1	STE18574	Y18574 Streptomyce
42	40.6	19.5	14210	1	AE012929	AE012929 Chlorobiu
43	40.4	19.4	4556	1	TFSCSAB	X54073 T.flavus sc
44	40.2	19.3	1242	6	E13276	E13276 Tamias asia
45	40.2	19.3	1367	10	AB000545	AB000545 Tamias s1

ALIGNMENTS

RESULT 1
MTU318818
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

MTU318818
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1415-97.
610 bp DNA circular BCT 09-AUG-2002

AJ318818
RNA polymerase beta subunit; rpoB gene.
Mycobacterium tuberculosis.
Mycobacterium tuberculosis
Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1
Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE Molecular analysis of rifampin-resistant *Mycobacterium tuberculosis* isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 610)

AUTHORS Herrera, L.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES Location/Qualifiers

source 1..610

gene /organism="Mycobacterium tuberculosis"

CDS /isolate="1415-97"

<1..>610 /db_xref="taxon:1773"

/gene="rpoB"

/codon_start=1

/transl_table=11

/product="RNA polymerase beta subunit"

/protein_id="CAC87035.1"

/db_xref="GI:22208413"

/translation="LDIYKRLRPEPTPKESAQTLLLENLFFKERYDLARVGRYKVNK KLGLHVGEPITSSLTEDVATIEYVRLHGGOTTMTVPGVEVPVETDDIDHFGNR RLRTVGEIIONOIRVGMSEMRVVRMTTQDVEAITPOTLINIRPVVAIKEFFGTS QPSQFMGQNNPLSGLTHKRLSALPGGLSRERAGLEVRDHP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 610;

Best Local Similarity 87.5%; Pred. No. 1.1e-21;

Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGCTAGGACCTGGCCGCTGTCGGCCGATACAAAGTCAACAAAGCTGG 60

Db 80 TCAAGGAGAACGCTAGGACCTGGCCGCTGTCGGCCGATACAAAGTCAACAAAGCTGG 139

QY 61 GCCTGAACACCAATCATCCGATCACCACGACGCTGACCGAAGAGACGCTCGCGCCA 120

Db 140 GGCTGCATGTCGGGAGCCCATCACGTCGACGCTGACCGAAGAGACGCTCGCGCCA 199

QY 121 CCATCAGATATCTGTCGGCTGCACGAGGCCACGATGACCGTGGCGGGGG 180

Db 200 CCATCGAATATCTGTCGGCTGCACGAGGTCAGACACGATGACCGTTCGGGGCGCG 259

QY 181 TCGAGTGGCGGTGGAACCGACGACAT 208

Db 260 TCGAGTGGCGGTGGAACCGACGACAT 287

RESULT 2

MTU318819

LOCUS

DEFINITION

610 bp DNA circular BCT 09-AUG-2002

Myobacterium tuberculosis partial rpoB gene for RNA polymerase beta subunit, isolate 1417-97.

ACCESSION

AJ318819

VERSION

GI:22208414

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

Myobacterium tuberculosis.

ORGANISM

Myobacterium tuberculosis.

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE

1

AUTHORS

Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE

Molecular analysis of rifampin-resistant *Mycobacterium tuberculosis* isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 610)

AUTHORS

Herrera, L.

TITLE Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES Location/Qualifiers

source 1..610

gene /organism="Mycobacterium tuberculosis"

CDS /isolate="1417-97"

<1..>610 /db_xref="taxon:1773"

/gene="rpoB"

/codon_start=1

/transl_table=11

/product="RNA polymerase beta subunit"

/protein_id="CAC87036.1"

/db_xref="GI:22208415"

/translation="LDIYKRLRPEPTPKESAQTLLLENLFFKERYDLARVGRYKVNK KLGLHVGEPITSSLTEDVATIEYVRLHGGOTTMTVPGVEVPVETDDIDHFGNR RLRTVGEIIONOIRVGMSEMRVVRMTTQDVEAITPOTLINIRPVVAIKEFFGTS QPSQFMGQNNPLSGLTHKRLSALPGGLSRERAGLEVRDHP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 610;

Best Local Similarity 87.5%; Pred. No. 1.1e-21;

Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACGCTAGGACCTGGCCGCTGTCGGCCGATACAAAGTCAACAAAGCTGG 60

Db 80 TCAAGGAGAACGCTAGGACCTGGCCGCTGTCGGCCGATACAAAGTCAACAAAGCTGG 139

QY 61 GCCTGAACACCAATCATCCGATCACCACGACGCTGACCGAAGAGACGCTCGCGCCA 120

Db 140 GGCTGCATGTCGGGAGCCCATCACGTCGACGCTGACCGAAGAGACGCTCGCGCCA 199

QY 121 CCATCAGATATCTGTCGGCTGCACGAGGCCACGATGACCGTGGCGGGGG 180

Db 200 CCATCGAATATCTGTCGGCTGCACGAGGTCAGACACGATGACCGTTCGGGGCGCG 259

QY 181 TCGAGTGGCGGTGGAACCGACGACAT 208

Db 260 TCGAGTGGCGGTGGAACCGACGACAT 287

RESULT 3

MTU318813

LOCUS

DEFINITION

616 bp DNA circular BCT 09-AUG-2002

Myobacterium tuberculosis partial rpoB gene for RNA polymerase beta subunit, isolate 1763-97.

ACCESSION

AJ318813

VERSION

GI:22208402

KEYWORDS

RNA polymerase beta subunit; rpoB gene.

SOURCE

Myobacterium tuberculosis.

ORGANISM

Myobacterium tuberculosis.

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE

1

AUTHORS

Herrera, L., Jimenez, M.S. and Saez, J.A.

TITLE

Molecular analysis of rifampin-resistant *Mycobacterium tuberculosis* isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 616)

AUTHORS

Herrera, L.

TITLE

Direct Submission

JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES Location/Qualifiers

source 1..616

/translation="TTEALLDIYRKLRPGEPTTKESAQTLLLENLFFKRYDLARVGR
YKNNKGLHVGEPITSTLTEDVATIEYLVRLHEGQTTMTVPGEVPEVETDDID
HFGNRLRTVGLIQNIIRVGMRSRMRVVRMTTQDVEAITPOTLINIRPVAAIKE
FFGTSPQSFQMGQNNPLSLGTHKRLSALPGGLSRERAGLEVRDV"
BASE COUNT 124 a 192 c 207 g 95 t

Query Match 80.0%; Score 166.4; DB 1; Length 618;
Best Local Similarity 87.5%; Pred. No. 1.1e-21;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
ORIGIN
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCCGATACAAAGTCAACAAGAGCTGG 60
Db 95 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCCGATACAAAGTCAACAAGAGCTGG 154
QY 61 GCCTGAACACCAATATCCGATCACACGAGCGCTGACCGAAGAAGACGCTGTCGCCCA 120
Db 155 GGCTGCATGTGGCGAGCCCATACGCTGTCGACGCTGACCGAAGAAGACGCTGTCGCCCA 214
QY 121 CCATCGAGTATCTGTCGCCCTGCACGAGGCCAGGCGCCACGATGACCGTGGCGGGGG 180
Db 215 CCATCGAATATCTGTCGCCCTGTCACGAGGCTGACAGCCACGATGACCGTTCGCGGGCGG 274
QY 181 TCGAGGTGCGGTGGAAACCGACGACAT 208
Db 275 TCGAGGTGCGGTGGAAACCGACGACAT 302

RESULT 6
MTU318814
LOCUS
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
ACCESSION AJ318814
VERSION AJ318814.1 GI:22208404
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis.
REFERENCE 1
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 633)
AUTHORS Herrera, L.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid. 28220, SPAIN

FEATURES
source
1..633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1..633
/gene="rpoB"
<1..>633
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87031.1"
/db_xref="GI:22208405"

BASE COUNT 129 a 195 c 210 g 99 t
ORIGIN
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCCGATACAAAGTCAACAAGAGCTGG 60

Query Match 80.0%; Score 166.4; DB 1; Length 633;
Best Local Similarity 87.5%; Pred. No. 1.1e-21;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
ORIGIN
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCCGATACAAAGTCAACAAGAGCTGG 60
Db 110 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCCGATACAAAGTCAACAAGAGCTGG 169
QY 61 GCCTGAACACCAATATCCGATCACACGAGCGCTGACCGAAGAAGACGCTGTCGCCCA 120
Db 170 GGCTGCATGTGGCGAGCCCATACGCTGTCGACGCTGACCGAAGAAGACGCTGTCGCCCA 229
QY 121 CCATCGAGTATCTGTCGCCCTGCACGAGGCCAGGCGCCACGATGACCGTGGCGGGGG 180
Db 230 CCATCGAATATCTGTCGCCCTGTCACGAGGCTGACAGCCACGATGACCGTTCGCGGGCGG 289
QY 181 TCGAGGTGCGGTGGAAACCGACGACAT 208
Db 290 TCGAGGTGCGGTGGAAACCGACGACAT 317

RESULT 7
MTU318816
LOCUS
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
ACCESSION AJ318816
VERSION AJ318816.1 GI:22208408
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis.
REFERENCE 1
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 637)
AUTHORS Herrera, L.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda, Madrid. 28220, SPAIN

FEATURES
source
1..637
/organism="Mycobacterium tuberculosis"
/isolate="1255-98"
/db_xref="taxon:1773"
1..637
/gene="rpoB"
<1..>637
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87033.1"
/db_xref="GI:22208409"

BASE COUNT 128 a 198 c 210 g 101 t
ORIGIN
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCCGATACAAAGTCAACAAGAGCTGG 60

Query Match 80.0%; Score 166.4; DB 1; Length 637;
Best Local Similarity 87.5%; Pred. No. 1.1e-21;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
ORIGIN
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGCCGATACAAAGTCAACAAGAGCTGG 60

```

|||||
Db 101 TCAAGGAGAGCGCTACGACCTGGCCCGCTCGTTCGCTATAAGTCAACAAGAAGCTCG 160
|||||
QY 61 GCCTGAACACCAATCATCGATCACACGACGACCTGACCGAAGAAGCTGTCGCGCA 120
|||||
Db 161 GCGTGCATGTCGGCGAGGCCATACGCTGCTGACGCTGACCGAAGAAGACGTCGTCGCCA 220
|||||
QY 121 CCATCGAGTATCTGCTCGCCCTGCACGAGGGCCAGGCCACGATGACCGTGCCTGGCGGGG 180
|||||
Db 221 CCATCGAATATCTGCTCGCTTGCACGAGGTCAGACCCAGCATGACCGTTCGGGGCGCG 280
|||||
QY 181 TCGAGGTCCCGTGGAAACCGAGCAT 208
|||||
Db 281 TCGAGGTCCCGTGGAAACCGAGCAT 308
|||||

RESULT 8
LOCUS MTU318821 639 bp DNA circular BCT 09-AUG-2002
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1071-98.
ACCESSION AJ318821
VERSION AJ318821.1 GI:22208418
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 639)
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5.,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
Source Location/Qualifiers
1..639
/organism="Mycobacterium tuberculosis"
/isolate="1071-98"
/db_xref="taxon:1773"
gene 1..639
CDS <1..>639
/gene="rpoB"
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87038.1"
/db_xref="GI:22208419"
/translation="RRIDEALDIYRKLRPGEPTKESQTLLENLFFKRYDLARV
GRYKVKLGHVGEPTSTLSEDVATIEYVLRHGEQTMVPGVEVPVETDD
IDHFGNRLRLTQELIQIRVMSRMRVRMTQDVEAITPTLLINIRPAVAAI
KEFGTSLSQFMVQNNPLSLGTLTKRRLSALGPGLSRERAGLEVDRVHPHS"
BASE COUNT 126 a 202 c 212 g 99 t
ORIGIN
Query Match 80.0%; Score 166.4; DB 1; Length 639;
Best Local Similarity 87.5%; Pred. No. 1.le-21;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCTCGTTCGCTATAAGTCAACAAGAAGCTCG 160
|||||
Db 101 TCAAGGAGAGCGCTACGACCTGGCCCGCTCGTTCGCTATAAGTCAACAAGAAGCTCG 160
|||||
QY 61 GCCTGAACACCAATCATCGATCACACGACGACCTGACCGAAGAAGCTGTCGCGCA 120
|||||
Db 161 GCGTGCATGTCGGCGAGGCCATACGCTGCTGACGCTGACCGAAGAAGACGTCGTCGCCA 220
|||||
QY 121 CCATCGAGTATCTGCTCGCCCTGCACGAGGGCCAGGCCACGATGACCGTGCCTGGCGGGG 180
|||||
Db 221 CCATCGAATATCTGCTCGCTTGCACGAGGTCAGACCCAGCATGACCGTTCGGGGCGCG 280
|||||
QY 181 TCGAGGTCCCGTGGAAACCGAGCAT 208
|||||
Db 281 TCGAGGTCCCGTGGAAACCGAGCAT 308
|||||

RESULT 9
LOCUS I50706 970 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 1 from patent US 5643723.
ACCESSION I50706
VERSION I50706.1 GI:2472409
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 970)
AUTHORS Persing, D.H., Hunt, J.J., Young, K.K.Y., Felmlee, T.A., Roberts, G.D.
and Whelan, A.Christian.
TITLE Detection of a genetic locus encoding resistance to rifampin in
mycobacterial cultures and in clinical specimens
JOURNAL Patent: US 5643723-A 1 01-JUL-1997;
FEATURES Location/Qualifiers
Source 1..970
/organism="unknown"
BASE COUNT 182 a 302 c 330 g 156 t
ORIGIN
Query Match 80.0%; Score 166.4; DB 6; Length 970;
Best Local Similarity 87.5%; Pred. No. 1.le-21;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCTCGTTCGCGCGATACAGGTCACACAGAAGCTCG 60
|||||
Db 26 TCAAGGAGAGCGCTACGACCTGGCCCGCTCGTTCGCGCGATACAGGTCACACAGAAGCTCG 85
|||||
QY 61 GCCTGAACACCAATCATCGATCACACGACGACCTGACCGAAGAAGACGTCGTCGCCA 120
|||||
Db 86 GCGTGCATGTCGGCGAGGCCATCACGTCGTCGACCGTACCGAAGAAGACGTCGTCGCCA 145
|||||
QY 121 CCATCGAGTATCTGCTCGCCCTGCACGAGGGCCAGGCCACGATGACCGTGCCTGGCGGGG 180
|||||
Db 146 CCATCGAATATCTGCTCGCTTGCACGAGGTCAGACCCAGCATGACCGTTCGGGGCGCG 205
|||||
QY 181 TCGAGGTCCCGTGGAAACCGAGCAT 208
|||||
Db 206 TCGAGGTCCCGTGGAAACCGAGCAT 233
|||||

RESULT 10
LOCUS AX111339 3534 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 2072 from Patent WO0123604.
ACCESSION AX111339
VERSION AX111339.1 GI:13927631
KEYWORDS
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis.
REFERENCE Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 3534)
AUTHORS Bergeron, M.G., Boissinot, M., Huletsky, A., m Nard, C., Ouellette, M.,
Picard, F.J. and Roy, P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: WO 0123604-A 2072 05-APR-2001;
FEATURES Infectio Diagnostic (I.D.I.) INC. (CA)
Location/Qualifiers

```

source	1. .3534	/organism="Mycobacterium tuberculosis"	
		/strain="Rv"	
BASE COUNT	679 a	/db_xref="taxon:1773"	
ORIGIN	1188 g	586 t	
Query Match	80.0%;	Score 166.4;	DB 6; Length 3534;
Best Local Similarity	87.5%;	Pred. No. 8.9e-22;	
Matches 182;	Conservative	0; Mismatches 26;	Indels 0; Gaps 0;
QY	1	TCAAGGAGAGCGCTACGACCTGCGCCGCTGCGCGGATACAAAGTCAACAAGAGCTGG	60
Db	902	TCAAGGAGAGCGCTACGACCTGCGCCGCTGCGCTGCTATAGGTCAACAAGAGCTGG	961
QY	61	GCCTGAACCAATCATCCGATCCACGACGACGCTGACCGAAGAAGACGTCGTCGCCA	120
Db	962	GGCTGCTATGTCGGCGAGCCCATCACCTGCTGACGCTGACCGAAGAAGACGTCGTCGCCA	1021
QY	121	CCATCGAGTATCTGTCGCGCTGCAGGAGCGCCAGGCGACGATGACCGTGCCTGGCGGGG	180
Db	1022	CCATCGAATATCTGTCGCGCTGCAGGAGGTCAGACCATGATGACCGTTCCTGGCGGGG	1081
QY	181	TCGAGTGCCTGGTGAACCGACGACAT	208
Db	1082	TCGAGTGCCTGGTGAACCGACGACAT	1109
RESULT 11			
MTU12205		3853 bp	DNA linear BCT 02-MAR-2000
LOCUS			
DEFINITION			Mycobacterium tuberculosis H37Rv RNA-polymerase beta subunit (rpoB) gene, partial cds.
ACCESSION			U12205
VERSION			U12205.1 GI:515684
KEYWORDS			Mycobacterium tuberculosis.
SOURCE			Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex. .
ORGANISM			1 (bases 1 to 3853)
REFERENCE			Imboden,P., Trollier,R., Marchesi,F., Telenti,A., Bodmer,T., Cole,S., Schopfer,K. and Burkart,T.
AUTHORS			The rpoB gene of Mycobacterium tuberculosis
TITLE			Unpublished
JOURNAL			2 (bases 1 to 3853)
REFERENCE			Imboden,P.
AUTHORS			Direct Submission
TITLE			Submitted (11-JUL-1994) Paul Imboden, Institute for Medical Microbiology, University of Berne, Friedbuehlstrasse 51, Berne, 3010, Switzerland
JOURNAL			
FEATURES			Location/Qualifiers
source			1. .3853
			/organism="Mycobacterium tuberculosis"
			/strain="H37Rv"
gene			/db_xref="taxon:1773"
			576. .>3853
			/gene="rpoB"
CDS			576. .>3853
			/gene="rpoB"
			/codon_start=1
			/transl_table=11
			/product="RNA-polymerase beta subunit"
			/protein_id="AAA20242.2"
			/db_xref="GI:7144493"
			/translation="MLEGCILADRSQSKTAASPSRQSSNNNSVPGAPNRYVFAKL REPLEVGLLDVQDSFEWLTGSRWRSAERGDVNPVGLLEEVLYEYSPIDFSGS MSLSFSDPRFDDKAPVDECKDKDMTYAAPLFVTAEP INNNTGEIKSOTVFMGDFPMW TEKGTFTINGRVVVSQVLRSPGVFDET IDKTDKTLHSAKV IPSRGAWLEFDYDK RDTGVGRIDRRKRPQVTVLLKALGWTSEQIVRGPSEINRSTLEKDNVGTGTEALLD IYKRLRGEPTKESATLLENLFFKEKRLDLARVGRYKVNKLGLHVGEPITTSLT EEDVATVIEYLVRUHEGQTTMTVPVGVEVPVETDDIDHFNRRRLRTVGLIQNIRVG

MSRMRVVRERMTTQDVEAITPQTLINIRPVAAAIKEFFGT SOLSFMDONNPLSGIT HKRRLSALGPGLSRERAGLEVDRVHPSHYGRMCPTETPEGNIGLIGSLSVYRVNP FGFETPYRKVVDVSDGIVYLTADEEDRHVQAQNSPIDADGRVFRVLRKAG EYEVSPSEVDYMDVSPRONVSATAMIPLEHDDANRALMGANMQAVPLVRSAP LVSTGHELRADIAADATYSSQESGVIEVSADYITVMDNGTRRTYRMRKPARNSHGT ANQCPIVDAGDRVAGQVIADGPTDDGEMALGNLLVAIMPMEHNYEDAITLSNRL VEEDLVTSIHIEHEIDARDTLKGAETIRDIPNIDEVLAIDLDERGIVRIGAEVRDG DILVGVTPKGETELTPEERLLRAIFGEKAREVRDTSKVPHGESCVMIDGTPVDI EDELPAGVNPRMNYVIAQKRKISDGLAGRHGNGVIGKILPVEDMFLADGTPVDI ILNTHGVPRMNIGOIILEHLCHCAHSGHKVDAAGKVPDWAARLPDELLEAQNALYS TPFVDAQAEALAGLSCTILPNRDGDVLDADKAMLFDRSGEPFPYPTVGYMTIM KLHLVDDKIHARSTGPYSMTIPLGGKAQFGGQRFGEHMECHMAQYGAAYTLOELL TIKS"					723 a	1173 c	1293 g	664 t
BASE COUNT								
ORIGIN								
Query Match 80.0%; Score 166.4; DB 1; Length 3853; Best Local Similarity 87.5%; Pred. No. 8.8e-22; Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;								
QY	1	TCAAGGAGAGCGCTACGACCTGCGCCGCTGTCGGCCGATACAAGGTCAACAAGAGCTGG	60					
Db	1477	TCAAGGAGAGCGCTACGACCTGCGCCGCTGCTGCTATAGGTCAACAAGAGCTCG	1536					
QY	61	GCCTGAACCAATCATCCGATCCACGACGACGCTGACCGAAGAAGACGTCGTCGCCA	120					
Db	1537	GGCTGCTATGTCGGCGAGCCCATCACCTGCTGCGCTGCTATAGGTCAACAAGAGCTCG	1596					
QY	121	CCATCGAGTATCTGTCGCGCTGCAGGAGCGCCAGGCGACGATGACCGTGCCTGGCGGG	180					
Db	1597	CCATCGAATATCTGTCGCGCTTGCACGAGGTCAGACCATGATGACCGTTCCTGGCGGG	1656					
QY	181	TCGAGGTGCCGTGGAACCGACGACAT	208					
Db	1657	TCGAGGTGCCGTGGAACCGACGACAT	1684					
RESULT 12								
MSGRPOB 5084 bp DNA linear BCT 13-SEP-1994								
LOCUS								
DEFINITION Mycobacterium tuberculosis RNA polymerase beta-subunit (rpoB) gene, complete cds and RNA polymerase beta'-subunit rpoC gene, partial cds.								
L27989								
L27989.1 GI:468333								
RNA polymerase beta-subunit; rpoB gene.								
Mycobacterium tuberculosis (strain Rv) DNA.								
Mycobacterium tuberculosis								
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;								
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.								
1 (bases 1 to 5084)								
Miller,L.P., Crawford,J.T. and Shinnick,T.M.								
The rpoB gene of Mycobacterium tuberculosis								
Antimicrob. Agents Chemother. 38 (4), 805-811 (1994)								
94304130								
PUBMED 8031050								
FEATURES								
source								
1. .5084								
/organism="Mycobacterium tuberculosis"								
/strain="Rv"								
/db_xref="taxon:1773"								
1065. .4598								
/gene="rpoB"								
1065. .4598								
/gene="rpoB"								
/codon_start=1								
/evidence=experimental								
/transl_table=11								
/product="RNA polymerase beta-subunit"								
/protein_id="AAA21416.1"								
/db_xref="GI:468334"								
/translation="MLEGCILADRSQSKTAASPSRQSSNNNSVPGAPNRYVFAKL"								

REFERENCE
AUTHORS

1 (bases 1 to 19770)
Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,
Harris, D., Gordon, S.V., Eigmeier, K., Gas, S., Barry III, C.E.,
Tekaiia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,
Connor, R., Davies, K., Devlin, K., Feltwell, T., Gentles, S.,
Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J.,
Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M.A.,
Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, S.,
Squares, S., Squires, R., Sulston, J.E., Taylor, K., Whitehead, S. and
Barrell, B.G.
Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence
Nature 393 (6685), 537-544 (1998)
98295987
9634230
2 (bases 1 to 19770)
Parkhill, J.
Direct Submission
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
tuberculosis sequencing and mapping teams. Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 27, 1998 this sequence version replaced gi:2143285.

COMMENT

Notes:
Details of M. tuberculosis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have
been renumbered from the original cosmid submissions but the old
gene designations are in brackets after the new gene numbers.
Gene prediction was based on a Hidden Markov Model of TB genes
implemented in TBparse (Krogh) supplemented with visual inspection
of positional base preference in codons, especially where there is
an increase in the observed/expected third position G + C.
CAUTION: In some cases we may not have predicted the correct
initiation codon. Where possible we choose an initiation codon
(atg, gtg, or ttg) which is preceded by an upstream ribosome
binding site sequence (optimally 5-13bp before the initiation
codon). If this cannot be identified we choose the most upstream
initiation codon.

FEATURES
source

Location/Qualifiers
1. .19770
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
/clone="1376"
57. .61
/note="aggA, possible rbs upstream of Rv0654"
68. .1573
/gene="Rv0654"
68. .1573
/gene="Rv0654"
/note="Rv0654", (MTCI376.22), len: 501. unknown, FASTA
score: Q53353 LIGNOSTILBENE-ALPHA, BETA-DIOXYGENASE (485
aa) opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity
in 523 aa overlap). Also similar to M. tuberculosis protein
MTCY212.07c (29.5% identity in 522 aa overlap)."
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0654"
/protein_id="CAB09380.1"
/db_xref="GI:2143307"
/db_xref="SPTREMBL:O06785"
/translation="MTQAQAESQNPYLEGLFAPVSTVATDLPVTGRIPHELDGRY
LNGPNPAEVDPAHYHPTGDMVGVALRDKGKRWYRNKVRPAPCAALGEFISA
RHPRTGIIEGGPNVNLTHAGTILALVEAGVYNYELDELDTVGCPDFDGLHGHT
APDQRPDGHGHNVSYSFARGHRYQVSYIGTDGHARTVDIEVAGSPMHHSFSLTDN
YVYIDLPVTFDPMQVPASVPVRLQRPALVQSVLGRVTRIPDIAALGNRMVCHSD
RLPYANWPSYPARGVMPREGNEDRWFDIEPCVYHPLNAYSRCNGAEVILVDV
RYSRMFDRRCGGDSRPSLDRTWTLNLTAGVTAECRDDRAQEPRIENLYGGPHR
FAYTVGIEGFLVGAGALSTPLYLKQDCVTGSGTSVQSLDPLDIGEWFVFPNSARAE
DGIILMGVHGRGREGQLLLDLDLAOTLESIAIVHLQVPVPMGFHGNWAPT"
1585. .2664

gene

CDS

/gene="Rv0655"
1585. .2664
/gene="Rv0655"
/note="Rv0655", (MTCI376.21), len: 359. abc transporter,
FASTA score: YRBF_ECOLI_P45393 hypothetical abc
transporter atp-binding (269 aa) opt: 644 z-score: 721.8
E(): 3.4e-33 (38.5% identity in 244 aa overlap); contains
PS00017 ATP/GTP-binding site motif A, PS00211 ABC
transporters family signature, highly similar to M. leprae
MKL_MYCLE_P30769 possible ribonucleotide transport atp-
(347 aa) opt: 2021 z-score:2244.4 E(): 0, (92.2% identity
in 335 aa overlap). Also similar to many other M.
tuberculosis ABC transporters eg. MTCY253.24 (33.6%
identity in 241 aa overlap)."
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0655"
/protein_id="CAB09379.1"
/db_xref="GI:2143306"
/db_xref="SWISS-PROT:O06784"
/translation="MRYSDSYHTTGRMOPRASTGEFPMGVSIEVNGLTKSFGSSRIWE
DVTLTIPAGEVSVLLGPSTGKSVFLKSLIGLLRPERGSIIDGTDIIECSAKELYEI
RTLGVLPQDQALFGSMNLXDNTAFPLREHTKKKESEIRDIYMEKALVGLGGDEKFF
PGEISGGNRKAGLARALVLDPQIILCDEPDGLDPVRTAYLSQLINDINAOIDATIL
IVYTHNIARTVPDNNMGLFRKHLVMEFGPREVLLTSDEPVVRQFLNGRRIPGIMSEE
KDEATMAEEQALLDAGHHAGGVEETGVPVPOISATPGMPERKAVARQARVREMLHTL
PKKAAAILDDLEGTHKYAVHEIQ"
1762. .1785
/gene="Rv0655"
/note="PS00017 ATP/GTP-binding site motif A"
2074. .2118
/gene="Rv0655"
/note="PS00211 ABC transporters family signature"
complement(3052. .3435)
/gene="Rv0656c"
complement(3052. .3435)
/gene="Rv0656c"
/note="Rv0656c", (MTCI376.20, unknown), len: 127 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0656c"
/protein_id="CAB09378.1"
/db_xref="GI:2143305"
/db_xref="SPTREMBL:O06783"
/translation="MAAATTTGTHRGLELRAAQRAVGSCEPQRAEFCRSARNADFFDQ
MSRMFGDVPDPVPKPSVYMRIDTSIAQHLRAGAVGALSVVVDLLICDTAAAGLUWLH
DDADYELAERHLPDIRVRVVSADD"
complement(3530. .3685)
/gene="Rv0657c"
complement(3530. .3685)
/gene="Rv0657c"
/note="Rv0657c", (MTCI376.19), unknown, len: 51 aa; similar
to several other M. tuberculosis hypothetical proteins eg.
YW08_MYCTU_Q10848 hypothetical 8.9 kd protein cy39.08c (80
aa) fasta scores; opt: 107 z-score: 182.3 E(): 0.0038,
45.8% identity in 48 aa overlap. Also similar to MTCY48_5
and AL020958|SC4H8_7 Streptomyces coelicolor cosmid 4H8 (66
aa), 41.0% identity in 39 aa overlap."
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0657c"
/protein_id="CAB09377.1"
/db_xref="GI:2143304"
/db_xref="SPTREMBL:O06782"
/translation="MSVTQIDLDELADVNRITAAVHTKKKAVNLAMRDYVERFRRIE
ALARSRE"
complement(3691. .3696)
/note="possible RBS upstream of Rv0657c"
complement(3761. .4477)
/gene="Rv0658c"
complement(3761. .4477)
/gene="Rv0658c"
/note="Rv0658c", (MTCI376.18), len: 238, unknown,

Db 1248 TCGAGTTCGCGTCGAGTCGCACGACAT 1275

Search completed: November 13, 2002, 01:25:56
Job time : 662.723 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:59:42 ; Search time 1141.13 Seconds
(without alignments)
3037.202 Million cell updates/sec

Title: US-09-697-123B-7
Perfect score: 214
Sequence: 1 tcaagggaagcgctacgac.....ccggtcgagtcgacgacat 214

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum Match 100%
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
EST:*			
1:	em_estba:*		
2:	em_esthum:*		
3:	em_estin:*		
4:	em_estmu:*		
5:	em_estov:*		
6:	em_estpl:*		
7:	em_estro:*		
8:	em_htc:*		
9:	gb_est1:*		
10:	gb_est2:*		
11:	gb_htc:*		
12:	gb_est3:*		
13:	gb_est4:*		
14:	gb_est5:*		
15:	em_estfun:*		
16:	em_estom:*		
17:	gb_gss:*		
18:	em_gss_hum:*		
19:	em_gss_inv:*		
20:	em_gss_pln:*		
21:	em_gss_vrt:*		
22:	em_gss_fun:*		
23:	em_gss_mam:*		
24:	em_gss_mus:*		
25:	em_gss_other:*		
26:	em_gss_pro:*		
27:	em_gss_rod:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
c 1	139.2	65.0	1282 9	AI770311 42 Mycoba
c 2	47.4	22.1	1850 11	AY107051
c 3	45	21.0	569 17	BH897054
4	44.6	20.8	665 13	BM601475
5	44.6	20.8	687 13	BM590640
6	43.8	20.5	275 14	R86542

7	43.6	20.4	371	10	AV438805
8	43.2	20.2	461	14	C73000
9	43	20.1	234	14	R86588
10	43	20.1	235	14	R86566
11	43	20.1	331	14	R86510
12	43	20.1	878	12	BF264852
13	42.6	19.9	618	13	BJ270031
14	42.6	19.9	623	13	BJ270031
15	42.4	19.8	479	10	BE361556
16	42.2	19.7	279	14	R86528
17	42.2	19.7	710	12	BG321153
18	42	19.6	2598	11	AY103647
19	41.6	19.4	1856	11	AY109418
20	41.4	19.3	416	12	BF484291
21	41.4	19.3	714	14	BQ842534
22	40.8	19.1	1006	12	BG106122
23	40.8	19.1	1750	11	BC029105
24	40.6	19.0	434	13	BI578713
c 25	40	18.7	257	17	BH804652
c 26	39.8	18.6	407	10	BE406021
27	39.8	18.6	525	10	AV929492
28	39.8	18.6	550	13	BJ208458
29	39.8	18.6	558	12	BE754521
30	39.8	18.6	579	13	BJ208491
31	39.8	18.6	585	17	BH806894
32	39.8	18.6	589	13	BJ244734
33	39.8	18.6	613	13	BJ260431
34	39.8	18.6	624	13	BJ248032
35	39.8	18.6	633	13	BJ317529
36	39.8	18.6	645	10	AW056150
37	39.8	18.6	645	14	BQ241492
38	39.8	18.6	672	13	BJ244763
39	39.8	18.6	680	14	BQ820533
40	39.8	18.6	792	12	BF627944
41	39.8	18.6	840	10	AW448084
42	39.8	18.6	840	10	BQ606019
43	39.8	18.6	925	17	CNS0091P
44	39.6	18.5	417	10	BE363166
45	39.6	18.5	500	10	BE360701

ALIGNMENTS

RESULT 1
AI770311/c 1282 bp mRNA linear EST 24-JAN-2000
LOCUS 42 Mycobacterium anaerobic stationary phase library Mycobacterium
DEFINITION smegmatis cDNA, mRNA sequence.
ACCESSION AI770311 GI:6742680
VERSION AI770311.1
KEYWORDS EST.
SOURCE Mycobacterium smegmatis.
ORGANISM Mycobacterium smegmatis.
REFERENCE 1 (bases 1 to 1282)
AUTHORS Murugasu-Oei, B., Tay, A. and Dick, T.
TITLE Upregulation of stress response genes and ABC transporters in
an aerobic stationary-phase Mycobacterium smegmatis
JOURNAL Mol. Gen. Genet. 262 (4-5), 677-682 (1999)
MEDLINE 20092472
COMMENT Contact: Murugasu-Oei, B.
Mycobacterium Laboratory
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore 117609, Republic of Singapore
Tel: 65 874 3011
Fax: 65 779 1117
Email: mcbome@imcb.nus.edu.sg
Insert Length: 1282 Std Error: 0.00
Seq primer: T3 Forward: T7 Backward.

FEATURES
source location/Qualifiers
1. .1282

AV438805	AV438805
C73000	C73000
R86588	R86588
R86566	R86566
R86510	R86510
BF264852	BF264852
BJ270031	BJ270031
BJ270031	BJ270031
BE361556	BE361556
DG1_72_E0	DG1_72_E0
R86528	R86528
RG321153	RG321153
2m04_0510	2m04_0510
AY103647	AY103647
AY109418	AY109418
zeta may	zeta may
BF484291	BF484291
WHE2321_E	WHE2321_E
BQ842534	BQ842534
WHE2994_B	WHE2994_B
BG106122	BG106122
602289945	602289945
BC029105	BC029105
Homo sapi	Homo sapi
BI578713	BI578713
RE72617_5	RE72617_5
BH804652	BH804652
1008097B0	1008097B0
WHE0405_f	WHE0405_f
AV929492	AV929492
BJ208458	BJ208458
BE754521	BE754521
208060_MA	208060_MA
BJ208491	BJ208491
BH806894	BH806894
1008064A0	1008064A0
BJ260431	BJ260431
BJ244734	BJ244734
BJ248032	BJ248032
BJ317529	BJ317529
AW056150	AW056150
660004G09	660004G09
BQ241492	BQ241492
TAE05004D	TAE05004D
BJ244763	BJ244763
BQ820533	BQ820533
1030085E0	1030085E0
BF627944	BF627944
HVSMEB000	HVSMEB000
AW448084	AW448084
BRY_1624	BRY_1624
BQ606019	BQ606019
BRY_1624	BRY_1624
AL053013	AL053013
Drosophila	Drosophila
BE363166	BE363166
DG1_9_F12	DG1_9_F12
BE360701	BE360701
DG1_66_D0	DG1_66_D0

```

/organism="Mycobacterium smegmatis"
/strain="mc2155"
/db_xref="taxon:1772"
/clone_lib="Mycobacterium anaerobic stationary phase
library"
/lab_host="E. coli XL1-Blue MRF"
/notes="Vector: Lambda ZAP II; Bacilli were disrupted using
a Mini bead beater (Biospec). RNA was isolated using the
RNeasy protocol (Qiagen). Purified RNA was subjected to 2
rounds of digestion with RNase-free DNase I (Promega).
DNase I was heat-inactivated at 75oC for 5 min. and
removed by using RNeasy columns followed by phenol
extraction and ethanol precipitation. The RNA
preparations were confirmed to be free of genomic DNA
contamination by carrying out PCR and RT-PCR using the
Access kit (Promega) and primers specific for the
histone-like protein gene hlp (Lee et al., 1998). cDNA
was synthesized using random hexamer primers (Promega)
and Stratagene's cDNA synthesis kit. cDNA fragments were
ligated into lambda ZAP II vector and packaged in vitro
using Stratagene reagents."
BASE COUNT      230 a 395 c 405 g 252 t
ORIGIN

Query Match      65.0%; Score 139.2; DB 9; Length 1282;
Best Local Similarity 81.8%; Pred. No. 7.4e-21;
Matches 175; Conservative 0; Mismatches 33; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGTGGCGGTACAAAGTGAACAAGAGCTGG 60
Db 236 TCAAGGAGAGCGCTACGACCTGGCCGCGTGGCGGTACAAAGTGAACAAGAGCTGG 177

QY 61 GTCTTGGCGGTGCCAACCCGCGTCTGGTGACTGCCACACGCTCACCGAGGAGACGTCG 120
Db 176 GCCTGAACGGGCAAGCC-----GATCACCAGCTCGACGCTGACCGAAGAGGACGTCG 123

QY 121 TCCCCACCATCGGTACCTGGTGGCGCTGCACGAGGGCCAGACACGATGACCCGCCCGCG 180
Db 122 TCCGACCATCGAGTACCTGGTGGCGTCTGCACGAGGGTCAGACTTCGATGACCGCTCCCGG 63

QY 181 GCGGCTCGAGGTCGCGGTTCGAGGTCGACGACAT 214
Db 62 GTGGCTCGAGGTTCCCGTCGAGGTCGACGACAT 29

RESULT 2
AY107051
LOCUS      AY107051      1850 bp      mRNA      linear      HTC 25-MAY-2002
DEFINITION Zea mays PCO135033 mRNA sequence.
ACCESSION  AY107051
VERSION     AY107051.1  GI:21210129
KEYWORDS   HTC.
SOURCE     Zea mays.
ORGANISM   Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
REFERENCE  2 (bases 1 to 1850)
AUTHORS   Coe,E.C.
TITLE     Direct Submission
JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
LOCATION/Qualifiers
1. .1850
/organism="Zea mays"
/db_xref="MaizeDB:637903"
/db_xref="taxon:4577"

```

```

/clone="PCO135033"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
BASE COUNT      331 a 625 c 583 g 311 t
ORIGIN

Query Match      22.1%; Score 47.4; DB 11; Length 1850;
Best Local Similarity 52.2%; Pred. No. 0.96;
Matches 105; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 12 CGCTACGACCTGGCCGCGTGGCGGTGTAACAAGAGCTGGGTGGCGGT 71
Db 260 CGGTCTACCAAGAGTGGCGCGGTACGGCCCATCATGACGGTGTGGCTGGCGACG 319

QY 72 GCCAACCCGCTCTGGTGACTGCCACACGCTCACCGAGGAGACGTCGTCGCCACCATC 131
Db 320 TCGCCACGGTGGTGTGTCCACGTCGGAGCTGCCAAGGAGGTGCTCAAGACCCACGAC 379

QY 132 GGTACCTGGTGGCGCTGCACGAGGGCCAGACACGATGACCCGCCCGCGCTCGAG 191
Db 380 CAGCAGCTCGCGACGCGTGGCGGCGCTCCACCGAGAGCTTCAGCCGAGGCGGCAG 439

QY 192 GTCCCGGTGCGAGGTCGACGAC 212
Db 440 GACCTCATCTGGCGCGACTAC 460

RESULT 3
BH897054/c
LOCUS      BH897054      569 bp      DNA      linear      GSS 14-AUG-2002
DEFINITION 3526.1_6.1_E11.1EL_Y_1 3526 - RescueMu Grid K Zea mays genomic, DNA
sequence.
ACCESSION  BH897054
VERSION     BH897054.1  GI:22232454
KEYWORDS   GSS.
SOURCE     Zea mays.
ORGANISM   Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
Walbot,V.
Maize genomic sequences found using engineered RescueMu transposon
1 (bases 1 to 569)
Unpublished (2001)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Very probable ligation site of ends cut by single endonuclease.
Reverse complemented post-ligation sequence from source sequence.
Plate: 3526.1_6.1 row: 6
Class: transposon-tagged.
LOCATION/Qualifiers
1. .569
/organism="Zea mays"
/cultivar="mixed background W23/A188/B73"
/db_xref="taxon:4577"
/clone_lib="3526 - RescueMu Grid K"
/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Organ: leaf; Vector: RescueMu (engineered from
pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
RescueMu is a 4.9 kb, modified maize Mu transposon

```



```

Matches 104: Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 8 GAAGCCCTACGACCTGGCCCGGTGGCGGTACAAAGTGAACAAAGACCTGGTCTTGG 67
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 255 GCAGCGGACGAGTGGTGGCGCTGATGGCGCGCAACTCCGACGACGTGGCGCGTCTGCT 314
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 68 CGGTGCCAACCGCGCTCTGGTACTGCCACACGCTACCCGAGGAGAGCTGCTGCCAC 127
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 315 CGTTCCTTCCTGGCGGTGTAACTGTCAGCAGCTCGATCGCTGCTTCGGCGTGAGGA 374
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 128 CATCGGGTACCTGGTGGCTGACAGAGGCGGACACGACGATGACGCCCGCGGCT 187
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 375 GGTGGAGCACCTGCTGCTGCGCTACCCGCGCGCAACGTGATAGCGGATGCGGACGCT 434
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 188 CGAGGTCCCGGTGCGAGTCCGACG 210
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435 GCCGTGCTCTACGAAGACCG 457
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
R86542
LOCUS
DEFINITION
R86542 RABEST103T Rabbit Osteoclast, Dennis Sakai Oryctolagus cuniculus
CDNA clone PRABOC103 5' similar to cofillin, mRNA sequence.
ACCESSION
R86542
VERSION
R86542.1 GI:947196
KEYWORDS
EST.
SOURCE
rabbit.
ORGANISM
Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE
1 (bases 1 to 275)
AUTHORS
Sakai,D., Tong,H.-S. and Minkin,C.
TITLE
Osteoclast Molecular Phenotyping by Random cDNA Sequencing
JOURNAL
Bone 17 (2), 111-119 (1995)
MEDLINE
96021365
COMMENT
Other ESTs: RABEST038T, RABEST045T, RABEST084T, RABEST140T,
RABEST163T, RABEST202T, RABEST229T
Contact: Sakai D
Basic Sciences
University of Southern California
USC School of Dentistry, 925 West 34th Street, DEN-4220, Los
Angeles, CA 90089-0641
Tel: 213/7405563
Fax: 213/7407560
Email: sakai@molbio.usc.edu
Seq primer: T7 promoter.
FEATURES
Location/Qualifiers
source
1..275
/organism="Oryctolagus cuniculus"
/strain="New Zealand White"
/db_xref="taxon:9986"
/clone="PRABOC103"
/clone_lib="Rabbit Osteoclast, Dennis Sakai"
/lab_host="E. coli DH12S"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; Poly(A)
)+ RNA was purified from a 97% pure population of
osteoclasts prepared from the long bones of 10 day old
rabbits. First strand cDNA was synthesized by priming
with an oligo(dT)-NotI anchor-primer and second strand
cDNA was synthesized by replacement synthesis as described
by Gubler and Hoffman (Gene 25:283, 1983). Following the
addition of SalI adapters and NotI digestion, the cDNA was
cloned between the SalI (5o) and NotI (3o) sites of the
pSPORT1 (BRL) plasmid vector."
BASE COUNT
67 a 76 c 92 g 40 t
ORIGIN

Query Match 20.5%; Score 43.8; DB 14; Length 275;
Best Local Similarity 53.1; Pred. No. 4.6;
Matches 93; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 5 GGAGAAGCGCTACGACCTGGCGCGGTGGCGGTACAAAGTGAAGAGCTGGTCT 64
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 13 GAAGAAGCGCAAGAAGCGGTGCTGTCTGCTGAGCGAGGACAAGAAGACATCGTCT 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 65 TGGCGGTGCCAACCGCGCTCTGGTACTGCCACACGCTCACCAGGAGACGCTCGTCCG 124
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 73 GGAGAGGGCAAGAGATCTCTGGTGGCGAGCTGGGGCCAGACCGTGGAGACCCCTACGC 132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 125 CACCATCGGTACTGCTGTGCGCTTGCACGAGGCGCAGACACGATGACCCGCCGCC 179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 133 CACCTTCGTGAGATGCTGCCGACAAGGACTGCCGCTACGCCCTCTTACGACGCC 187
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
AV438805
LOCUS
DEFINITION
AV438805 Porphyra yezoensis TU-1 Porphyra yezoensis cDNA clone
PS041e12_r 5', mRNA sequence.
ACCESSION
AV438805
VERSION
AV438805.1 GI:8594030
KEYWORDS
EST.
SOURCE
Porphyra yezoensis.
ORGANISM
Porphyra yezoensis.
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.
REFERENCE
1 (bases 1 to 371)
AUTHORS
Nikaido,I., Asamizu,E., Nakajima,M., Nakamura,Y., Saga,N. and
Tabata,S.
TITLE
Generation of 10,154 expressed sequence tags from a leafy
gametophyte of a marine red alga, Porphyra yezoensis
JOURNAL
DNA Res. 7, 223-227 (2000)
MEDLINE
20363100
COMMENT
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp. URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
Location/Qualifiers
source
1..371
/organism="Porphyra yezoensis"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PS041e12_r"
/clone_lib="Porphyra yezoensis TU-1"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT
40 a 140 c 139 g 52 t
ORIGIN

Query Match 20.4%; Score 43.6; DB 10; Length 371;
Best Local Similarity 56.2%; Pred. No. 5.3;
Matches 82; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 67 GCGGTGCCAACCGCGCTCTGGTACTGCCACACGCTCACCAGGAGACGCTCGTCCGCA 126
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 24 GCGCGGCGCCCAACCGGTGGTGACCGCCACACCGCGCGCGCTCACCGCCA 83
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 127 CCATCGGTACTGTTGGCGCTTGCACGAGGCGCAGACACGATGACCCGCCGCCGCC 186
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 84 GCGCGGCGCACCTGGTGTATCAATGGCGCGCGCGCGCTGCGGTGCGGGTGG 143
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 187 TCGAGTCCCGGTGCGAGGTGCGAGCAG 212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 144 GCGACCGCTGGACGTGGCGCGCGCGAC 169
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
C73000
LOCUS
DEFINITION
C73000 Rice panicle at flowering stage Oryza sativa (Japonica
cultivar-group) cDNA clone E2649_1A, mRNA sequence.
ACCESSION
C73000
VERSION
C73000.1 GI:2428537
KEYWORDS
EST.

```

[illegible]

/clone="pRBOC140"
 /clone_lib="Rabbit Osteoclast, Dennis Sakai"
 /lab_host="E. coli DH12s"
 /notes="vector: pSPORT1; Site_1: SalI; Site_2: NotI; Poly(A)
) + RNA was purified from a 97% pure population of
 osteoclasts prepared from the long bones of 10 day old
 rabbits. First strand cDNA was synthesized by priming
 with an oligo(dT)-NotI anchor-primer and second strand
 cDNA was synthesized by replacement synthesis as described
 by Gubler and Hoffman (Gene 25:283, 1983). Following the
 addition of SalI adapters and NotI digestion, the cDNA was
 cloned between the SalI (5o) and NotI (3o) sites of the
 pSPORT1 (BRL) plasmid vector."
 57 a 63 c 81 q 34 t.

[illegible]

RESULT	11
LOCUS	R86510
DEFINITION	RABEST202T Rabbit Osteoclast, Dennis Sakai Oryctolagus cuniculus CDNA clone PRABOC202 5' similar to cofillin, mRNA sequence.
ACCESSION	R86510
VERSION	R86510.1
KEYWORDS	GI:947262
SOURCE	EST. rabbit
ORGANISM	Oryctolagus cuniculus Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE	1 (bases 1 to 331)
AUTHORS	Sakai,D., Tong,H.-S. and Minkin,C.
TITLE	Osteoclast Molecular Phenotyping by Random cDNA Sequencing
JOURNAL	Bone 17 (2), 111-119 (1995)
MEDLINE	96021365
COMMENT	Other ESTs: RABEST03QT RABEST04ST RABEST04TM RABEST04CM

Contact: Sakai D
Basic Sciences
University of Southern California
USC School of Dentistry, 925 West 34th Street, Los
Angeles, CA 90089-0641
Tel: 2137405563
Fax: 2137407560
Email: sakai@molbio.usc.edu

```

FEATURES
source
seq primer: T/ promoter.
Location/Qualifiers
1. .331
/organism="Oryctolagus cuniculus"
/strain="New Zealand White"
/db_xref="taxon:9986"
/clone="PRABOC202"
/clone_lib="Rabbit Osteoclast, Dennis Sakai"
/lab_host="E. coli DH12S"
/note="Vector: pSP0T11; Site.1: SalI; Site.2: NotI; Poly(A)
)+ RNA was purified from a 97% pure population of
osteoclasts prepared from the long bones of 10 day old
rabbits. First strand cDNA was synthesized by priming

```

with an oligo(dT)-NotI anchor-primer and second strand cDNA was synthesized by replacement synthesis as described by Gubler and Hoffman (Gene 25:283, 1983). Following the addition of SalI adaptors and NotI digestion, the cDNA was cloned between the SalI (50') and NotI (30') sites of the pSPORT1 (BRL) plasmid vector."

Query Match 20.1%; Score 43; DB 14; Length 331;
Best Local Similarity 54.8%; Pred. NO. 7.1;
Matches 85; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

[illegible]

RESULT 12	
BF264852	
LOCUS	
DEFINITION	

LOCUS	BF264852	878 bp	mRNA	linear	EST 23-OCT-2001
DEFINITION	HV_CEA0010124f Hordeum vulgare seedling green leaf EST library				
	HVCDA0004 (Blumeria challenged) Hordeum vulgare cDNA clone				
	HV_CEA0010124f, mRNA sequence.				

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE

JOURNAL
COMMENT

On Nov 17, 2000 this sequence version replaced g1:11195846.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: wing@clemson.edu
Total hq bases = 224
Seq primer: AATTAACCTCACTAAAGGG
High quality sequence stop: 348.

FEATURES
source

```

1. 0760
/organism="Hordeum vulgare"
/cultivar="C116155 (M1a13)"
/db_xref="taxon:4513"
/clone="HV_CEA0010124F"
/clone_lib="Hordeum vulgare seedling green leaf EST
library HVCNANO0004 (Blumeria challenged)"
/tissue_type="seedling green leaf"
/lab_host="TJC121"
/notes="Vector: lambda2AP; Site_1: EcoRI; Site_2: XhoI;
C.1. 16155 (M1a13) plants were greenhouse grown in the R
Wise lab at Iowa State University, Ames, IA; 7 day old
green seedlings were challenged with isolate A27 (Avm1a13

```



```

/db_xref="taxon:4565"
/clone_lib="y. Ogihara unpublished cDNA library, wh_dL"
/clone_lib="y. Ogihara unpublished cDNA library, wh_dL"
/tissue_type="crown of seedling"
/dev_stage="freekes' scale 1"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid:
Site.1: EcoRI; Site.2: XhoI; Plants were grown under
hydroponic conditions at UC Davis, salt stressed for 12
hours, and for 7 days, then dissected and frozen (Akhunov
in J. Dvorak Lab). Total RNA was prepared from sheath
tissue, equal quantities of RNA were pooled from the two
samples, polyA was purified from the pooled RNA, a cDNA
library was made, and the cDNA clones were in vivo
excised to give pBluescript phagemids in the TJ Close lab
at the University of California, Riverside (Akhunov, Chin
, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).
Plasmid DNA preparations and DNA sequencing were
performed in the OD Anderson lab (all other authors)."
BASE COUNT      114 a      208 g      84 t
ORIGIN
Query Match      19.9%; Score 42.6; DB 13; Length 623;
Best Local Similarity 50.2%; Pred. No. 9.4;
Matches 105; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 2 CAAGGAGAACGGCTACGACCTGGCCGGCTGCGGTTACAAGGTGAACAAGAGCTGGG 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 CAACGCCGAGGTCTACGCCAACGGCGCGGAGGAGATCATGGGCGGCGATCCGGGA 273

QY 62 TCTTGGGGTGCCAACCCGGCTCTGTGTGACTGCCACCGCTCACCGAGGAGAGCTGGT 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 CTTGGGCTGGCGCGCGGAGGTGCTCATCTCCACCAAGCTCTTCTGGGCGGCGCAGG 333

QY 122 CGCACCATCGGTACTGTGGCTGCTGCACGAGGCGGACACGATGACCGCCCCCGG 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 CCCACGACAGGGGCTCTCCCGAAGCACATCTCGAGGGGCTCAAGGGCTCGCTCAA 393

QY 182 CGGCTCGAGGTCCCGTCTGAGTGTGACG 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 GCGGCTCAACATGGAGTACGTGACGTCG 422

```

```

RESULT 15
BE361556
LOCUS      DGI_72_E07.b1_A002 Dark Grown 1 (DGI) Sorghum bicolor cDNA, mRNA
DEFINITION sequence.
ACCESSION BE361556
VERSION   BE361556.1 GI:9303015
KEYWORDS  EST.
SOURCE    sorghum.
ORGANISM  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
          clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 479)
AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
          ,L.H.
          An EST database from Sorghum: dark-grown seedlings
          Unpublished (2000)
          Contact: Cordonnier-Pratt MM
          Laboratory for Genomics and Bioinformatics
          The University of Georgia, Department of Plant Biology
          Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
          Tel: 706 542 1860
          Fax: 706 583 0210
          Email: mmpratt@uga.edu
          Sequences have been trimmed to exclude PolyA, vector and regions
          below Phred quality 16. The threshold for highest quality sequence
          is 20.
          Seq primer: JEN REV
          High quality sequence stop: 456
          POLYA-No.

```

```

FEATURES
    source              Location/Qualifiers
    1..479
        /organism="Sorghum bicolor"
        /db_xref="taxon:4558"
        /clone_lib="Dark Grown 1 (DGI)"
        /note="Organ: 5-day-old dark-grown seedlings; Vector:
        Lambda Zap; Site.1: XhoI; Site.2: EcoRI; The library was
        made from poly-A RNA in the cloning vector lambda ZAP II.
        Clones to be sequenced were prepared by mass excision."
BASE COUNT      85 a     158 c     170 g     66 t
ORIGIN
Query Match      19.8%; Score 42.4; DB 10; Length 479;
Best Local Similarity 50.5%; Pred. No. 10;
Matches 103; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 11 GCGCTACGACCTGGCCCGCGTGGCGGTTACAAGGTGAACAAGAGCTGGTCTTTGGCGG 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 GCCCAAGGCCGTGGTGGAGGTGCTGCAAGACACGCGCATCAAGAAGGTGAAGCTTTCTGA 273

QY 71 TGCCAAACCCGGCTCTGTGTGACTGCGCACGCTCACCGAGGAGACGCTGCTCGCCACCAT 130
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 CACCGACCCGCGCCCATGAGCGCGCTCGCGCGGACCGGCATCGAGGTCTATGTCGCCCAT 333

QY 131 CGGGTACTGTGGCGCTGCGACGAGGCGCCAGACCATGACCGCCCCCGGCGGCTCGA 190
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 CCCCAACAACATGTCTCGCGGACCTCGCGGAGCGAGCGGAAAGGCTAAGGACTGGGTCAA 393

QY 191 GGTCCCGGTGAGGTGCGACGACAT 214
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 GCGCAAGCTCGCGGCTACGACTT 417

```

Search completed: November 13, 2002, 03:59:41
Job time : 1156.63 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 : Search time 21.7711 Seconds
(without alignments)
3487.380 Million cell updates/sec

Title: US-09-697-123B-7

Perfect score: 214

Sequence: 1 tcaaggagagcgtacgac.....ccggtcgaggtcgacgacat 214

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	DB ID	Description
1	88.4	41.3	5096	10	US-09-984-711-5	Sequence 5, Appl1
2	88.4	41.3	5099	9	US-10-075-460-5	Sequence 5, Appl1
3	88.4	41.3	5099	9	US-09-887-052-1	Sequence 1, Appl1
4	88.4	41.3	5099	10	US-09-887-052-3	Sequence 3, Appl1
5	88.4	41.3	5099	10	US-09-887-052-5	Sequence 5, Appl1
6	40	18.7	3343	10	US-09-950-772-7	Sequence 7, Appl1
7	40	18.7	15872	10	US-09-861-289-1	Sequence 1, Appl1
8	38	17.8	1380	10	US-09-950-772-3	Sequence 3, Appl1
9	37.6	17.6	2898	10	US-09-737-149-3	Sequence 3, Appl1
10	36.6	17.1	584	10	US-09-867-550-669	Sequence 669, App
11	36.2	16.9	221	10	US-09-864-761-26660	Sequence 26660, A
12	36.2	16.9	392	10	US-09-878-574-3598	Sequence 3598, App
13	36.2	16.9	1914	10	US-09-815-242-7960	Sequence 7960, Ap
14	36.2	16.9	3057	10	US-09-815-242-4131	Sequence 4131, Ap
15	35.6	16.6	242	10	US-09-923-876-2475	Sequence 2475, Ap
16	35.6	16.6	244	10	US-09-923-876-2527	Sequence 2527, Ap
17	35.6	16.6	248	10	US-09-923-876-1166	Sequence 1166, Ap
18	35.6	16.6	248	10	US-09-923-876-2335	Sequence 2335, Ap
19	35.6	16.6	260	10	US-09-923-876-2086	Sequence 2086, Ap

ALIGNMENTS

RESULT 1

US-09-984-711-5
; Sequence 5, Application US/09984711
; Patent No. US20020119549A1
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: BATHE, Brigitte
; APPLICANT: STEPHAN, Hans
; APPLICANT: KREUTZER, Caroline
; APPLICANT: HERMANN, Thomas
; APPLICANT: PFEFFERLE, Walter
; APPLICANT: BINDER, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE
; FILE REFERENCE: 204209US0
; CURRENT APPLICATION NUMBER: US/09/984,711
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: DE10108230.9
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 5096
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
; OTHER INFORMATION:
US-09-984-711-5

Query Match 41.3%; Score 88.4; DB 10; Length 5096;

Best Local Similarity 68.4%; Pred. No. 2.7e+13;

Matches 141; Conservative 0; Mismatches 56; Indels 9; Gaps 1;

QY 9 AAGCGGTACGACCTGCCCGCGTGGCGGTACAAAGTGAACAAAGCTGGGTCTTGGC 68
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1578 AAGCGGTACGACCTGCCCGTGGTGTGCGTTACAAGTCAACCGCAAGCTCGSCCTTGGT 1637

QY 69 GGTGCCAACCCGCTCTGGTGACTGCCACACCTCACCAGGAGAACGCTGCGCCACC 128
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 1638 GCGGACCAACGATGGTTTGATGACT-----CTTACTGAAGAGGACATCGCAACACC 1688

Sequence 2207, Ap
Sequence 3681, Ap
Sequence 5, Appl1
Sequence 7, Appl1
Sequence 2224, Ap
Sequence 2583, Ap
Sequence 7982, Ap
Sequence 7989, Ap
Sequence 7873, Ap
Sequence 1223, Ap
Sequence 3702, Ap
Sequence 1, Appl1
Sequence 2, Appl1
Sequence 3537, Ap
Sequence 320, App
Sequence 13, Appl1
Sequence 28400, A
Sequence 2448, Ap
Sequence 3118, Ap
Sequence 11830, A
Sequence 7875, Ap
Sequence 99, Appl1
Sequence 7772, Ap
Sequence 1, Appl1
Sequence 1695, Ap
Sequence 494, App

20 35.6 16.6 266 10 US-09-923-876-2207
21 35.6 16.6 312 10 US-09-923-876-3681
22 35.6 16.6 2916 10 US-09-737-149-5
23 35.6 16.6 3132 10 US-09-737-149-7
24 35.2 16.4 235 10 US-09-923-876-2224
25 35.2 16.4 254 10 US-09-923-876-2583
26 35.2 16.4 1434 10 US-09-815-242-7982
27 35.2 16.4 1818 10 US-09-815-242-7989
28 34.8 16.3 888 10 US-09-815-242-7873
c 29 34.6 16.2 446 10 US-09-815-343-1273
30 34.6 16.2 1642 10 US-09-880-107-3702
31 34.4 16.1 1539 10 US-09-732-680A-1
32 34.4 16.1 12606 10 US-09-957-974-2
33 34.2 16.0 239 10 US-09-923-876-3537
34 34.2 16.0 1713 10 US-09-954-456-320
c 35 34.2 16.0 2223 10 US-09-897-214-13
36 34 15.9 215 10 US-09-864-761-28400
37 34 15.9 224 10 US-09-923-876-2448
38 34 15.9 229 10 US-09-923-876-3118
39 34 15.9 468 10 US-09-864-761-11830
40 33.8 15.8 660 10 US-09-815-242-7875
c 41 33.8 15.8 1476 10 US-09-997-664-99
42 33.8 15.8 1992 10 US-09-815-242-7772
c 43 33.8 15.8 6491 10 US-09-997-664-1
44 33.6 15.7 258 10 US-09-923-876-1695
45 33.6 15.7 939 10 US-09-822-830A-494

```
QY 129 ATCGGTACCTGCTGGCTGCACGAGGCGCAGACCATGACCGCCCGCGGCGCTC 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1689 ATCGAGTACCTGCTGGCTGCACCGAGTGAGCCGCTGATGACTTCTCCAATGGTGAA 1748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 189 GAGGTCCCGGTGAGGTGACGACAT 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1749 GAGATCCAGTCGAGACCGATGACAT 1774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
US-10-075-460-5
; Sequence 5, Application US/10075460
; Patent No. US2002015557A1
; GENERAL INFORMATION:
; APPLICANT: MOCKEL, BETTINA
; APPLICANT: BATHE, BRIGITTE
; APPLICANT: HANS, STEFAN
; APPLICANT: KREUTZER, CAROLINE
; APPLICANT: HERMANN, THOMAS
; APPLICANT: PFEFFERLE, WALTER
; APPLICANT: BINDER, MICHAEL
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE
; FILE REFERENCE: 218472USOX
; CURRENT APPLICATION NUMBER: US/10/075,460
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: DE 10107230.9
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: DE 10162386.0
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
; OTHER INFORMATION:
US-10-075-460-5

Query Match 41.3%; Score 88.4; DB 9; Length 5099;
Best Local Similarity 68.4%; Pred. No. 2.7e-13;
Matches 141; Conservative 0; Mismatches 56; Indels 9; Gaps 1;

QY 9 AAGCGCTACGACCTGGCCGCGTGGCGGTACAAAGGTGAACAAGAGCTGGGTCTTGGC 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1578 AAGCGCTACGACCTGGCTGGCTGGTTCGTACAAAGTCAACCGCAAGCTCGGCCTTGGT 1637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 69 GGTGCCAACCCGGCTCTGGTGACTGCCACACGCTCACCAGAGGAAGACGTGCTGCGCAC 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1638 GCGCACACGATGGTTTGATGACT-----CTTACTGAAGAGGACATCGCAACCCACC 1688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 129 ATCGGTACCTGCTGGCTGCACGAGGCGCAGACCATGACCGCCCGCGGCGCTC 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1689 ATCGAGTACCTGCTGGCTGTCACGAGGTGAGCGCGTCTGCTCTCTCAAAATGGTGAA 1748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 189 GAGGTCCCGGTGAGGTGACGACAT 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1749 GAGATCCAGTCGAGACCGATGACAT 1774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-09-887-052-3
; Sequence 3, Application US/09887052
; Patent No. US20020119537A1
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: BATHE, Brigitte
; APPLICANT: HERMANN, Thomas
; APPLICANT: PFEFFERLE, Walter
; APPLICANT: BINDER, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE
; FILE REFERENCE: 204212USOX
; CURRENT APPLICATION NUMBER: US/09/887,052
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: DE1010729.5
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
; OTHER INFORMATION:
US-09-887-052-3

Query Match 41.3%; Score 88.4; DB 10; Length 5099;
Best Local Similarity 68.4%; Pred. No. 2.7e-13;
Matches 141; Conservative 0; Mismatches 56; Indels 9; Gaps 1;

QY 9 AAGCGCTACGACCTGGCCGCGTGGCGGTACAAAGGTGAACAAGAGCTGGGTCTTGGC 68
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1578 AAGCGCTACGACCTGGCTGGCTGGTTCGTACAAAGTCAACCGCAAGCTCGGCCTTGGT 1637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 69 GGTGCCAACCCGGCTCTGGTGACTGCCACACGCTCACCAGAGGAAGACGTGCTGCGCAC 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1638 GCGCACACGATGGTTTGATGACT-----CTTACTGAAGAGGACATCGCAACCCACC 1688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 129 ATCGGTACCTGCTGGCTGCACGAGGCGCAGACCATGACCGCCCGCGGCGCTC 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1689 ATCGAGTACCTGCTGGCTGTCACGAGGTGAGCGCGTCTGCTCTCTCAAAATGGTGAA 1748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 189 GAGGTCCCGGTGAGGTGACGACAT 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1749 GAGATCCAGTCGAGACCGATGACAT 1774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
US-09-887-052-1
; Sequence 1, Application US/09887052
; Patent No. US20020119537A1
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: BATHE, Brigitte
; APPLICANT: HERMANN, Thomas
; APPLICANT: PFEFFERLE, Walter
; APPLICANT: BINDER, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE
```

Db 1638 GCGCACCACGATGTTTGATGACT-----CTTACTGAAGAGACATCGCAACACC 1688
Qy 129 ATCGGGTACCTGTGGCCCTGCGACGAGGCGCCAGACACGATAGCAGCCCGCGGCGCTC 188
Db 1689 ATCAGGTACCTGGTGGCTCTGCGACGAGGTGAGCGGCTCATGACTTCTCCAAATGGTGA 1748
Qy 189 GAGGTCCCGGTGCGAGGTGCGAGCAT 214
Db 1749 GAGATCCCGAGTCGAGACCGATGACAT 1774

RESULT 5

US-09-887-052-5
; Sequence 5, Application US/09887052
; Patent No. US20020119537A1
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: BATHE, Brigitte
; APPLICANT: HERMANN, Thomas
; APPLICANT: PFEFFERLE, Walter
; APPLICANT: BINDER, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE
; FILE REFERENCE: 204212US0X
; CURRENT APPLICATION NUMBER: US/09/887,052
; PRIOR FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: DE10107229.5
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
US-09-887-052-5

Query Match 41.3%; Score 88.4; DB 10; Length 5099;
Best Local Similarity 68.4%; Pred. No. 2.7e-13;
Matches 141; Conservative 0; Mismatches 56; Indels 9; Gaps 1;
Qy 9 AAGCGCTACGACCTGGCGCGGTGCGGTGCAAGGTGAACAAGAGCTGGGTCTTTGGC 68
Db 1578 AAGCGCTACGACCTGGCTCGGTGCTGTACAAATCAACCGCAGCTCGGCCCTTGGT 1637
Qy 69 GGTGCCAACCGCGCTCTGGTACTGCCACACGCTCACCGAGGAAGACGTGCTGCCACC 128
Db 1638 GCGCACCACGATGTTTGATGACT-----CTTACTGAAGAGGACATCGCAACACC 1688
Qy 129 ATCGGGTACCTGTGGCCCTGCGACGAGGCGCCAGACACGATGACCGCCCGCGGCGCTC 188
Db 1689 ATCAGGTACCTGGTGGCTCTGCGACGAGGTGAGCGGCTCATGACTTCTCCAAATGGTGA 1748
Qy 189 GAGGTCCCGGTGCGAGGTGCGAGCAT 214
Db 1749 GAGATCCCGAGTCGAGACCGATGACAT 1774

RESULT 6

US-09-950-772-7
; Sequence 7, Application US/09950772
; Patent No. US20020102713A1
; GENERAL INFORMATION:
; APPLICANT: SUZUKI, Shunichi
; APPLICANT: ONISHI, No. US20020102713A1Imasa
; APPLICANT: YOKOZEKI, Kenzo
; TITLE OF INVENTION: 5-SUBSTITUTED HYDANTOIN RACEMASE, DNA CODING FOR THE RACEMASE, AND
; FILE REFERENCE: 212318US00
; CURRENT APPLICATION NUMBER: US/09/950,772
; CURRENT FILING DATE: 2001-09-13

; PRIOR APPLICATION NUMBER: JP2000-278571
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: JP2001-65815
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 3343
; TYPE: DNA
; ORGANISM: Microbacterium liquefaciens
US-09-950-772-7

Query Match 18.7%; Score 40; DB 10; Length 3343;
Best Local Similarity 51.1%; Pred. No. 0.13;
Matches 94; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
Qy 28 GCGTGGCGCGGTACAAAGGTGAACAAGAGCTGGGTCTTGGCGGTGCCAACCGGCTCTGG 87
Db 721 GCGATCTCATGTTTCGACGTCATTGTGAAGAACTGTCGAGTGTTCAGTCAAGGCGCATCA 780
Qy 88 TGACTGCCACACGCTCACCGAGGAGGAGAGCTGCTGCCACCATCGGGTACCTGGTGCGCC 147
Db 781 TCGAAGCCGACATCTCTGTGAAGGACGCGCGGATCGCGCCATCAGCGAGGAGCCCTCG 840
Qy 148 TGCACGAGGCGCCAGACACGATGACCGCCCGCGCGCTCGAGCTCCGGTTCGAGGTGCG 207
Db 841 AGCGCGAAGCCCGCGGACCATCATGATGCCGAGGAGTTCGTGATGCCCGGTGGTTCG 900
Qy 208 ACGA 211
Db 901 ATGA 904

RESULT 7

US-09-861-289-1
; Sequence 1, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438USI
; CURRENT APPLICATION NUMBER: US/09/861,289
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105,537
; PRIOR FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 15872
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-861-289-1

Query Match 18.7%; Score 40; DB 10; Length 15872;
Best Local Similarity 53.1%; Pred. No. 0.15;
Matches 85; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
Qy 53 GAAGCTGGGTCTTGGCGGTGCCAACCCGGCTCTGTGTGACTGCCACACCGCTCACCGAGGA 112
Db 2416 GGAGCTCGGTCCCGACGGGTCTGTCCCGATGCGCGGACTCCGTACGCGACGAGA 2475
Qy 113 AGACGTCTGTGCCACCATCGGGTACTGTGCGCTGCACAGGCGCCAGACACCATGAC 172
Db 2476 GCGCGCCACGGCGGTCTCGGCCCTCGGCAAGGCGCCCGGAGCCCAAGTCGCTGCTCGC 2535
Qy 173 CGCCCGCGCGGCGCTCGAGGTCCCGGTTCGAGTTCGACGAC 212
Db 2536 CGCACTCCACCGCTTCTCGTCCGGGCGCCACGACGTCGAC 2575

```
RESULT 8
US-09-950-772-3
; Sequence 3, Application US/09950772
; Patent No. US20020102713A1
; GENERAL INFORMATION:
; APPLICANT: SUZUKI, Shunichi
; APPLICANT: ONISHI, No. US20020102713Alimasa
; APPLICANT: YOKOZEKI, Kenzo
; TITLE OF INVENTION: 5-SUBSTITUTED HYDANTOIN RACEMASE, DNA CODING FOR THE RACEMASE, AND
; FILE REFERENCE: 212318USO
; CURRENT APPLICATION NUMBER: US/09/950,772
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: JP2000-278571
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: JP2001-65815
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Microbacterium liquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1377)
; OTHER INFORMATION:
US-09-950-772-3

Query Match      17.8%; Score 38; DB 10; Length 1380;
Best Local Similarity 51.1%; Pred. No. 0.38;
Matches 89; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 38 GTACAAGGTGAACAGAGCTGGTCTTTGGCGGTGCCAACCCGGCTCTGGTGACTGCCAC 97
DB 3 GTTCGACGCTATTGTGAAGAACTCTCGAGTGGTTTCCAGTCAGGCGCATCATCGAAGCCGA 62

QY 98 CACGCTCACCGAGGAGAGCTGCTGCCACCATCGGTACCTGGTGGCCCTGCACGAGGG 157
DB 63 CATCTCTGTGAAGACGCGCGGATCGCGCCATCAGGAGAGCCCTCGAGGCCGGAAGC 122

QY 158 CCAGCACACGATGACCCCGCGCGCTCGAGGTCCTCCGGTCTGAGGTCTGAGGTCTGACGA 211
DB 123 CGCCCGGACCATCGATCGCGCAGCAGGTTCGTGATGCCCGGCTGGTGTGATGA 176

RESULT 9
US-09-737-149-3
; Sequence 3, Application US/09737149
; Patent No. US20020077466A1
; GENERAL INFORMATION:
; APPLICANT: Spaderna, Steven K
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Muralidhara, Padigaru
; APPLICANT: Spytek, Kimberly A.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-620 CIP
; CURRENT APPLICATION NUMBER: US/09/737,149
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/170,564
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/173,165
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,362
; PRIOR FILING DATE: 1999-12-27
; PRIOR APPLICATION NUMBER: 60/173,544
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 60/174,404
; PRIOR FILING DATE: 2000-01-04
; PRIOR APPLICATION NUMBER: 60/174,962
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: 60/223,929
```

```
; PRIOR FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2898
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2898)
US-09-737-149-3

Query Match      17.6%; Score 37.6; DB 10; Length 2898;
Best Local Similarity 48.6%; Pred. No. 0.51;
Matches 103; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 2 CAAGGAGAAAGCGCTACGACCTGGCCGCGCTGGCGCGGTACAAGGTGAACAAGAAGCTGG 61
DB 1833 CAAGGGCTTCCGCTTCGGCACCGTGTGGAGAGACGCGCGGAGGTACATCAAGAAG 1892

QY 62 TCTTGGCGGTGCCAACCCGGCTCTGGTGACTGCCACACGCTCACCGAGGAAGAGCTGGT 121
DB 1893 CTTCCCGGACATGCACGCACACATGCGGCGCCACAGCGCGCCACACGCGCGCGGT 1952

QY 122 CGCCACCATCGGTACCTGGTGGCTGCACGAGGCGCCAGACCATGACCGCCCGGG 181
DB 1953 CGCCATGCTCACGAGCGACCCCGCCAAAGCTCAACGCTTTCATCATGGACAAGTCGCT 2012

QY 182 CGGCGCTCGAGGTCGCGGTGCGAGGTGCGACGACA 213
DB 2013 GGACTACGAGGTCTCCATCGACGCGGACTGCA 2044
```

```
RESULT 10
US-09-867-550-669
; Sequence 669, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells a
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 669
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (31)..(584)
; OTHER INFORMATION: wherein any n is one of a or t or g or c
US-09-867-550-669
```

```
Query Match      17.1%; Score 36.6; DB 10; Length 584;
Best Local Similarity 53.1%; Pred. No. 0.78;
Matches 78; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 68 CGGTGCAACCCGGCTCTGGTGACTGCCACACGCTCACCGAGGAGAGAGCTGTCGCCAC 127
DB 168 CAGCGGCTCGCGCTCGTCACCATCGCTGATCGCTGCGCGGCTCGGCGATCATGATGAC 227

QY 128 CATCGGTTACCTGGTGGCTGCGCTGCAGGAGGGCCAGACCATGATGACCGCGCGGGGCT 187
DB 228 CCTCGCGGACCTGAACGGGATCCACCGATCCGACATCTCGATGCCGAGGGGCGGTGCG 287
```

QY 188 CGAGGTCCTCCGGTTCGAGGTCGACGACAT 214
 Db 288 GCAGAACACCGGGCGAGGTCACATCCT 314

RESULT 11

```

US-09-864-761-26660
: Sequence 26660, Application US/09864761
: Patent NO. US2002048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED
: TITLE OF INVENTION: GENE EXPRESSION ANAL
: FILE REFERENCE: Aomica-x-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/006656
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine
: SEQ ID NO 26660
: LENGTH: 221

```

```

; LENGTH: 221
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; OTHER INFORMATION: MAP TO AL049569.13
; OTHER INFORMATION: EXPRESSED IN BT474. SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW. SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN. SIGNAL = 1
; OTHER INFORMATION: EST HIT: AB030176.1, EVALUATE 4.00e-97
; OTHER INFORMATION: EST HUMAN HIT: BC089551.1, EVALUATE 1.00e-94
; OTHER INFORMATION: SWISSPROT HIT: Q9Y2J8. EVALUATE 1.00e-29
; US-Q9-864-761-26660

```

Query Match 16.9%; Score 36.2; DB 10; Length 221;

[illegible]

RESULT 12

```

RESOLUTION 12
US-09-878-574-3598
; Sequence 3598, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 3598
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-008-Q1-B1-C6
US-09-878-574-3598

```

Query Match

	Query Match	16.94;	Score 36.2;	DB 10;	Length 392;
	Best Local Similarity	54.0%;	Pred. No. 0.95;		
	Matches 74;	Conservative 0;	Mismatches 63;	Indels 0;	Gaps 0;
QY	76	ACCCGGCTCTGTGACTGCCACACCGCTCACCGAGGAGACGTCGTGCCACCATCGGGT	135		
Db	2	ACCTGCCCAGGCGCTTCGGCTCATCGACCGGCAACACGCGGCGTCTACGCGGCCAGG	61		
QY	136	ACCTGTGTGGCTTGTCCACGAGGGCCAGACACGATGACGCCCCCGGCGGCTTCGAGGTCC	195		
Db	62	ACCTGAGGCCCTCTCTACGTGCTCTCGCGGGTGTCCGTCGCCCGACGACGTGGCGGTCA	121		
QY	196	CGGTGAGGTGCAGAC	212		
Db	122	TGCTGGCGGAGTCCAG	138		

RESULT 13

```

US-09-815-242-7960
; Sequence 7960, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242

```

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7960
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; NAME/KEY: CDS
; LOCATION: (1)...(1914)
US-09-815-242-7960

Query Match 16.9%; Score 36.2; DB 10; Length 1914;
Best Local Similarity 53.1%; Pred. No. 1.1;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 69 GGTGCCAACCCGGCTGTGTGACTGCCACACGCTCACCGAGGAAGAGCTGTGTGCCACC 128
   || || || || || || || || || || || || || || || || || || || || ||
Db 1615 GCGACGCGGTGTCCACGCGCCGCAAGATGATCACCGAGGCGGCGACAAAGCCACC 1674

QY 129 ATCGGTACTGTGTGGCGCTGCACGAGGCCACACCATGACCGCCCGCGGCGCTC 188
   || || || || || || || || || || || || || || || || || || || || ||
Db 1675 GCCGAGCAAGCGGCACCATCGAGAAGGCGCTGGGCGAGCTGGAAGCGCGGTGAAGGC 1734

QY 189 GAGTCTCCCGTCGAGTTCGACGACA 213
   || || || || || || || || || ||
Db 1735 GACGACAAGCGCGAGATCGAGGCCA 1759

RESULT 14
US-09-815-242-4131
; Sequence 4131, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; PROKARYOTES
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7960
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; NAME/KEY: CDS
; LOCATION: (1)...(1914)
US-09-815-242-7960

Query Match 16.9%; Score 36.2; DB 10; Length 1914;
Best Local Similarity 53.1%; Pred. No. 1.1;
Matches 77; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 69 GGTGCCAACCCGGCTGTGTGACTGCCACACGCTCACCGAGGAAGAGCTGTGTGCCACC 128
   || || || || || || || || || || || || || || || || || || || || ||
Db 1615 GCGACGCGGTGTCCACGCGCCGCAAGATGATCACCGAGGCGGCGACAAAGCCACC 1674

QY 129 ATCGGTACTGTGTGGCGCTGCACGAGGCCACACCATGACCGCCCGCGGCGCTC 188
   || || || || || || || || || || || || || || || || || || || || ||
Db 1675 GCCGAGCAAGCGGCACCATCGAGAAGGCGCTGGGCGAGCTGGAAGCGCGGTGAAGGC 1734

QY 189 GAGTCTCCCGTCGAGTTCGACGACA 213
   || || || || || || || || || ||
Db 1735 GACGACAAGCGCGAGATCGAGGCCA 1759

RESULT 14
US-09-815-242-4131
; Sequence 4131, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; PROKARYOTES
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7960
; LENGTH: 1914
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; NAME/KEY: CDS
; LOCATION: (1)...(1914)
US-09-815-242-7960
```

```

; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4131
; LENGTH: 3057
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-4131

Query Match 16.9%; Score 36.2; DB 10; Length 3057;
Best Local Similarity 55.0%; Pred. No. 1.1;
Matches 71; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 74 CAACCCGGCTGTGTGACTGCCACACGCTCACCGAGGAAGAGCTGTGTGCCACCACCG 133
   || || || || || || || || || || || || || || || || || || || || ||
Db 483 CCAGCCCAAGCTGCGACCCCTGCCCGGTATCGCCGAGGCGGAGATCTCTCGGCAACCGT 542

QY 134 GTACCTGGTGGCTGCACGAGGCCACACCATGATGACCGCCCGCGGCGCTCGAGGT 193
   || || || || || || || || || || || || || || || || || || || || ||
Db 543 GTTCGCCATGCGCCCTGTGTGACCCCGGTGAAGATGGCGCGTTCGCGCTCACCGCGG 602

QY 194 CCGCGTCA 202
   || || ||
Db 603 CGAGATCAA 611

RESULT 15
US-09-923-876-2475
; Sequence 2475, Application US/09923876
; Patent No. US20030013958A1
; GENERAL INFORMATION:
; APPLICANT: Laigudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0019-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 2475
; LENGTH: 2472
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700160665H1
US-09-923-876-2475

Query Match 16.6%; Score 35.6; DB 10; Length 242;
Best Local Similarity 57.0%; Pred. No. 1.3;
Matches 65; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 99 AGCTCACCGAGGAAGAGCTGTGTGCCACCATCGGTACCTGTGTGCGCTGCACGAGGC 158
   || || || || || || || || || || || || || || || || || || || || ||
Db 124 ATGTACCTGTGTGAGAGCCGCCGCCACACGAGCGGTCAAGACCCACCGTCTCGCCGCG 183

QY 159 CAGACACAGTACCGCCCGCGCGGCTTCAGGTCGCGGTTCGAGGTTCGAGGTTCGAGGAC 212
   || || || || || || || || || || || || || || || || || || || || ||
Db 184 CCGACCAAGGCCAGCGCGCGGCTTCGAGGCGGCCACCGAGGCGCGCGG 237

Search completed: November 12, 2002, 16:57:53
Job time : 38.7711 secs
```


Result No.	Score	Query	Length	DB	ID	Description
1	124.8	58.3	970	1	US-08-250-030-1	Sequence 1, Appli
2	124.8	58.3	970	5	PTC-US95-06790-1	Sequence 1, Appli
3	104	48.6	3447	2	US-08-313-185-57	Sequence 57, Appli
4	104	48.6	3447	3	US-09-082-614A-57	Sequence 1, Appli
5	43.2	20.2	4411529	4	US-09-103-840A-1	Sequence 10, Appli
6	41.2	19.3	1620	2	US-08-461-775-10	Sequence 10, Appli
7	41.2	19.3	1620	3	US-09-031-606-10	Sequence 10, Appli
8	41.2	19.3	2668	2	US-08-461-775-11	Sequence 11, Appli
9	41.2	19.3	2668	3	US-09-031-606-11	Sequence 11, Appli
10	41	19.2	1926	4	US-09-079-985-10	Sequence 10, Appli
11	41	19.2	3000	4	US-09-192-104-1	Sequence 1, Appli
12	41	19.2	3000	4	US-09-543-446-1	Sequence 1, Appli
13	40	18.7	1665	3	US-08-881-784-8	Sequence 8, Appli
14	40	18.7	1665	4	US-09-292-768-3	Sequence 3, Appli
15	40	18.7	1665	4	US-09-292-768-67	Sequence 67, Appli
16	40	18.7	1665	4	US-09-292-768-69	Sequence 69, Appli
17	40	18.7	15872	4	US-09-105-537-1	Sequence 1, Appli
18	39	18.2	1724	6	5352575-8	Patent No. 5352575
19	36.8	17.2	1950	3	US-08-911-853-14	Sequence 14, Appli
20	36.8	17.2	1950	4	US-09-479-409-14	Sequence 14, Appli
21	36.8	17.2	1950	4	US-09-479-453-14	Sequence 14, Appli
22	36.8	17.2	17612	3	US-08-911-853-29	Sequence 29, Appli
23	36.8	17.2	17612	4	US-09-479-409-29	Sequence 29, Appli
24	36.8	17.2	17612	4	US-09-479-453-29	Sequence 29, Appli
25	36.8	17.2	68750	3	US-09-335-409-1	Sequence 1, Appli
26	36.8	17.2	68750	4	US-09-568-102-1	Sequence 1, Appli
27	36.8	17.2	68750	4	US-09-567-969-1	Sequence 1, Appli


```

: Sequence 57, Application US/09082614A
: Patent No. 6124098
: GENERAL INFORMATION:
: APPLICANT: Heym, Beate
: APPLICANT: Cole, Stewart
: APPLICANT: Young, Douglas
: APPLICANT: Zhang, Ying
: APPLICANT: Honore, Nadine
: APPLICANT: Telenti, Amalio
: APPLICANT: Bodmer, Thomas
: TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
: TITLE OF INVENTION: In Mycobacterium Tuberculosis
: NUMBER OF SEQUENCES: 66
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSEE: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/082.614A
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/313,185
: FILING DATE: 12-OCT-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Meyers, Kenneth J.
: REGISTRATION NUMBER: 25,146
: REFERENCE/DOCKET NUMBER: 02356.0068-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 408-4000
: TELEFAX: (202) 408-4400
: INFORMATION FOR SEQ ID NO: 57:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3447 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-09-082-614A-57

Query Match 48.6%; Score 104; DB 3; Length 3447;
Best Local Similarity 71.5%; Pred. No. 1.8e-16;
Matches 153; Conservative 0; Mismatches 55; Indels 6; Gaps 1

Qy 1 TCAAGGAGAACGGCTACGACCTGGCCGCGTGGGGCCGTACAAGGTGAACAAGACGTGG 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 809 TCAAGGAGAAACGCTACGACCTGCCAGGTTGGTCGTTACAAGGTCAACAAGAGCTCG 868
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 61 GTCTTTGGCGGTGCCAACCCGCTCTGTTGACTGCCACACGCTCACCGAGGAAGACGTGG 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 869 GGTTCACGCGCGGTGAGT-----TGATCAGCTCGTCCAGCTGACCGAAGAGGATGTCG 922
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 121 TCGCCACCATCGGTGTTGCTGGCTCGACGAGGCGCCAGACCATGACCGGCCCGCG 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 923 TCGCCACCATAGAGTACCTGTTGCTCTGCATGAGGTCAGTCGCACAATGACTGTCCCG 982
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 181 GCGGCTCGAGTCCCGTTCGAGTTCGACGACAT 214
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 983 GTGGGTAGAGTGCCAGTGGAACTGACGATAT 1016
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
US-09-103-840A-1/c
: Sequence 1, Application US/09103840A

```

```
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 010830-035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1620 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1620
; US-08-461-775-10

Query Match          19.3%; Score 41.2; DB 2; Length 1620;
Best Local Similarity 49.5%; Pred. No. 0.12;
Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACCGCTACGACCTGGCCGCGTGGCGGTACAAGGTGAACAAGAAGCTGG 60
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1268 TCAAGGTCCTGGAGCAGAACCTCGCGCCACCGCGCAGCAGGACCGCGGTGCGCGGTGG 1327
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 61 GTCCTGGCGGTGCCAACCCCGCTCTGTGACTGCCACCGCTCACCGAGGAGAGAGCTGG 120
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1328 TCCGCGCGCGCGCTCGAGCCGCTCGGCTGGATCGCCGAGAACCGCGGCTCGAGGGCT 1387
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 121 TCGCCACCATCGGTTACCTGGTGGCTGCACGAGGCGCCAGACGATGACCGCCCGG 180
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1388 AGTCATCACCAACAGGTGGCGAGCTGCAAGGCGCGGCTTCAACGCGGCCACCG 1447
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 181 GCGGCTTCGAGGTCCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 214
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1448 GCGAGTACGCGACCTGGTCAAGCGCGCGGTCTCAT 1481

RESULT 7
US-09-031-606-10
; Sequence 10, Application US/09031606
; Patent No. 6153404
; GENERAL INFORMATION:
; APPLICANT: MAZODIER, Philippe
; APPLICANT: GUGLIEMI, Gerard
; TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
; TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031.606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/050.313
; FILING DATE: 10-MAY-1993
; APPLICATION NUMBER: FR 9011186
; FILING DATE: 10-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; TELECOMMUNICATION INFORMATION:

; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 010830-035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1620 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1620
; US-08-461-775-11

Query Match          19.3%; Score 41.2; DB 3; Length 1620;
Best Local Similarity 49.5%; Pred. No. 0.12;
Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 1 TCAAGGAGAACCGCTACGACCTGGCCGCGTGGCGGTACAAGGTGAACAAGAAGCTGG 60
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1368 TCAAGGTCCTGGAGCAGAACCTCGCGCCACCGCGCAGCAGGACCGCGGTGCGCGGTGG 1327
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 61 GTCCTGGCGGTGCCAACCCCGCTCTGTGACTGCCACCGCTCACCGAGGAGAGAGCTGG 120
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1328 TCCGCGCGCGCGCTCGAGCCGCTCGGCTGGATCGCCGAGAACCGCGGCTCGAGGGCT 1387
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 121 TCGCCACCATCGGTTACCTGGTGGCTGCACGAGGCGCCAGACGATGACCGCCCGG 180
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1388 AGTCATCACCAACAGGTGGCGAGCTGCAAGGCGCGGCTTCAACGCGGCCACCG 1447
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 181 GCGGCTTCGAGGTCCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 214
    |||||  |  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1448 GCGAGTACGCGACCTGGTCAAGCGCGCGGTCTCAT 1481

RESULT 8
US-08-461-775-11
; Sequence 11, Application US/08461775
; Patent No. 5858773
; GENERAL INFORMATION:
; APPLICANT: MAZODIER, Philippe
; APPLICANT: GUGLIEMI, Gerard
; TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
; TITLE OF INVENTION: INITIATION OF TRANSCRIPTION
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: George Mason Bldg., Washington & Prince Sts.
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461.775
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/050.313
; FILING DATE: 10-MAY-1993
; APPLICATION NUMBER: FR 9011186
; FILING DATE: 10-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 010830-035
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2668 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-461-775-11

Query Match 19.3%; Score 41.2; DB 2; Length 2668;
Best Local Similarity 49.5%; Pred. No. 0.13;
Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 1 TCAAGGAGAAGCGCTACGACCTGGCCCGCTGGCCCGGTACAAAGTGAACAAGAGCTGG 60
Db 2116 TCAAGGTCTTGACGACAACTTCGCGCGCACCGCGGACGAGGCCACCGGTGTGCGCGTCG 2175
QY 61 GTCTTGGGGTGCCAAACCCGGCTCTGTGACTGCCACACGCTCACCGAGGAAGACGTGG 120
Db 2176 TCCGCGCGCGCGCTCGAGCGCTGGCTGGATCGCGGAGAACGCCGCGCTCGAGGGCT 2235
QY 121 TCGCCACCATCGGTACCTGTGGCCCTGACGAGGGCCAGACACGATACCGCCCCCG 180
Db 2236 ACCTCATCACCAAGGTGGCGAGCTCGACAAGGGCCAGGGCTTCAACGCGGCCACCG 2295
QY 181 CGCGGCTCGAGTCCCGCTCGAGTCCGACGACAT 214
Db 2296 GCGAGTACGGGACCTGGTCAAGCGCGCGTCAT 2329

RESULT 9
US-09-031-606-11
Sequence 11, Application US/09031606
Patent No. 6153404
GENERAL INFORMATION:
APPLICANT: MAZODIER, Philippe
APPLICANT: GUGLIEMI, Gerard
TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
INITIATION OF TRANSCRIPTION
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,606
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/050,313
FILING DATE: 10-MAY-1993
APPLICATION NUMBER: FR 9011186
FILING DATE: 10-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 010830-035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-8620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 2668 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-031-606-11
Query Match 19.3%; Score 41.2; DB 3; Length 2668;
Best Local Similarity 49.5%; Pred. No. 0.13;
Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 1 TCAAGGAGAAGCGCTACGACCTGGCCCGCTGGCCCGGTACAAAGTGAACAAGAGCTGG 60
Db 2116 TCAAGGTCTTGACGACAACTTCGCGCGCACCGCGGACGAGGCCACCGGTGTGCGCGTCG 2175
QY 61 GTCTTGGGGTGCCAAACCCGGCTCTGTGACTGCCACACGCTCACCGAGGAAGACGTGG 120
Db 2176 TCCGCGCGCGCGCTCGAGCGCTGGCTGGATCGCGGAGAACGCCGCGCTCGAGGGCT 2235
QY 121 TCGCCACCATCGGTACCTGTGGCCCTGACGAGGGCCAGACACGATACCGCCCCCG 180
Db 2236 ACCTCATCACCAAGGTGGCGAGCTCGACAAGGGCCAGGGCTTCAACGCGGCCACCG 2295
QY 181 CGCGGCTCGAGTCCCGCTCGAGTCCGACGACAT 214
Db 2296 GCGAGTACGGGACCTGGTCAAGCGCGCGTCAT 2329

RESULT 10
US-09-079-955-10
Sequence 10, Application US/09079955A
Patent No. 6465209
GENERAL INFORMATION:
APPLICANT: Alexander Blinkovsky
APPLICANT: Kimberly Brown
APPLICANT: Elizabeth Gollightly
APPLICANT: Tony Byun
APPLICANT: Thomas Mathiasen
APPLICANT: Lene V. Kofod
APPLICANT: Mikio Fujii
APPLICANT: Chigusa Shizuoka
TITLE OF INVENTION: Methods For Producing Protein
FILE REFERENCE: 5253 500-US
CURRENT APPLICATION NUMBER: US/09/079,955A
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 10
LENGTH: 1926
TYPE: DNA
ORGANISM: Sphingomonas capsulata
US-09-079-955-10

Query Match 19.2%; Score 41; DB 4; Length 1926;
Best Local Similarity 51.4%; Pred. No. 0.14;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
QY 27 CGCGTGGCGCGGTACAAAGTGAACAAGCTGGGTCTTTGGCGGTGCCAACCCGGCTCTG 86
Db 1132 CGCGACTGGCGCTCGGTGGAAGACACGACGATGGATCCCATCTTCGCGCGCCGCAAGCCC 1191
QY 87 GTGACTGCCACACGCTCACCGAGGAAGAGCTGTCGCCACCATCGGTGTACCTGGTGGCG 146
Db 1192 AAGCCCTATTCTTCGCTTACCCGTTAACGAGGACTATTACACCGAAGGCGCGTGTGTGG 1251
QY 147 CTCGACGAGGGCCAGACGACGATGACGCGCCCGCGGCGCTCGAGGTCCCGGTCTCGAGTGC 206
Db 1252 CTGGAAGCGGACCAAGATCATCCCGGATGGCAGCGGCGGCAAGAGGGCTTGATGATTTC 1311
QY 207 GACCA 211
Db 1312 GCCAA 1316

```
RESULT 11
US-09-192-104-1
; Sequence 1, Application US/09192104B
; Patent No. 6184020
; GENERAL INFORMATION:
; APPLICANT: Alexander Blinkovsky
; APPLICANT: Tony Byun
; APPLICANT: Alan V. Klotz
; APPLICANT: Alan Sloma
; APPLICANT: Maria Tang
; APPLICANT: Mikio Fujii
; APPLICANT: Chigusa Marumoto
; APPLICANT: Lene Venke Kofod
; TITLE OF INVENTION: Polypeptides Having Amino-peptidase
; FILE REFERENCE: 5379,200-US
; CURRENT APPLICATION NUMBER: US/09/192,104B
; CURRENT FILING DATE: 1998-11-13
; EARLIER APPLICATION NUMBER: 60/069719
; EARLIER FILING DATE: 1997-12-16
; EARLIER APPLICATION NUMBER: 1465/97
; EARLIER FILING DATE: 1997-12-16
; EARLIER APPLICATION NUMBER: PA 1998 00670
; EARLIER FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3000
; TYPE: DNA
; ORGANISM: Sphingomonas
US-09-192-104-1

Query Match      19.2%; Score 41; DB 4; Length 3000;
Best Local Similarity 51.4%; Pred. No. 0.14;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 27 CGCGTGGCGCGGTACAAAGGTGAACAAAGAGTGGGTCTTTGGCGGTGCCAACCCGGCTCTG 86
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1801 CGCGACTGGCGCTCGGTGGAAGACACGACGATGGATCCATCTTCCGCCGCCGCAAGCCC 1860

QY 87 GTGACTGCCACCACTCAGCGAGAGACGCTCGTCCGCCACCATCGGGTACCTGGTGGC 146
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1861 AAGCCCTATTCTCGCTTACCGGTAAACGAGGACTATTACACCGAAGCGCGCTGGTGTG 1920

QY 147 CTGCAGAGGCGGACACCATGATGACCGCCCGCGCGCTCGAGGTCCCGGTGAGGTC 206
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1921 CTGGAAGCGGACCATGATCCCGGTAAACGAGGACTATTACACCGAAGCGCGCTGGTGTG 1980

QY 207 GACGA 211
    | | |
Db 1981 GCCAA 1985

RESULT 12
US-09-543-446-1
; Sequence 1, Application US/09543446
; Patent No. 6303560
; GENERAL INFORMATION:
; APPLICANT: Alexander Blinkovsky
; APPLICANT: Tony Byun
; APPLICANT: Alan V. Klotz
; APPLICANT: Alan Sloma
; APPLICANT: Maria Tang
; APPLICANT: Mikio Fujii
; APPLICANT: Chigusa Marumoto
; APPLICANT: Lene Venke Kofod
; TITLE OF INVENTION: Polypeptides Having Amino-peptidase
; FILE REFERENCE: 5379,210-US
; CURRENT APPLICATION NUMBER: US/09/543,446
; CURRENT FILING DATE: 2000-04-05
; EARLIER APPLICATION NUMBER: 60/069719
; EARLIER FILING DATE: 1997-12-16
; EARLIER APPLICATION NUMBER: PA 1998 00670
; EARLIER FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3000
; TYPE: DNA
; ORGANISM: Sphingomonas
US-09-192-104-1

Query Match      19.2%; Score 41; DB 4; Length 3000;
Best Local Similarity 51.4%; Pred. No. 0.14;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 27 CGCGTGGCGCGGTACAAAGGTGAACAAAGAGTGGGTCTTTGGCGGTGCCAACCCGGCTCTG 86
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1801 CGCGACTGGCGCTCGGTGGAAGACACGACGATGGATCCATCTTCCGCCGCCGCAAGCCC 1860

QY 87 GTGACTGCCACCACTCAGCGAGAGACGCTCGTCCGCCACCATCGGGTACCTGGTGGC 146
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1861 AAGCCCTATTCTCGCTTACCGGTAAACGAGGACTATTACACCGAAGCGCGCTGGTGTG 1920

QY 147 CTGCAGAGGCGGACACCATGATGACCGCCCGCGCGCTCGAGGTCCCGGTGAGGTC 206
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1921 CTGGAAGCGGACCATGATCCCGGTAAACGAGGACTATTACACCGAAGCGCGCTGGTGTG 1980

QY 207 GACGA 211
    | | |
Db 1981 GCCAA 1985

RESULT 13
US-08-881-784-8
; Sequence 8, Application US/08881784
; Patent No. 6083731
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Lupien, Shari L.
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
; THE PRODUCTION OF LIMONENE HYDROXYLASES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
; ADDRESS: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,784
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelton, Dennis K.
; REGISTRATION NUMBER: 26,997
; REFERENCE/DOCKET NUMBER: WSUR19777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 224-0718
; TELEFAX: (206) 224-0779
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
```

```
; EARLIER FILING DATE: 1997-12-16
; EARLIER APPLICATION NUMBER: 1465/97
; EARLIER FILING DATE: 1997-12-16
; EARLIER APPLICATION NUMBER: PA 1998 00670
; EARLIER FILING DATE: 1998-05-15
; EARLIER APPLICATION NUMBER: 09/192,104
; EARLIER FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3000
; TYPE: DNA
; ORGANISM: Sphingomonas
US-09-543-446-1

Query Match      19.2%; Score 41; DB 4; Length 3000;
Best Local Similarity 51.4%; Pred. No. 0.14;
Matches 95; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 27 CGCGTGGCGCGGTACAAAGGTGAACAAAGAGTGGGTCTTTGGCGGTGCCAACCCGGCTCTG 86
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1801 CGCGACTGGCGCTCGGTGGAAGACACGACGATGGATCCATCTTCCGCCGCCGCAAGCCC 1860

QY 87 GTGACTGCCACCACTCAGCGAGAGACGCTCGTCCGCCACCATCGGGTACCTGGTGGC 146
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1861 AAGCCCTATTCTCGCTTACCGGTAAACGAGGACTATTACACCGAAGCGCGCTGGTGTG 1920

QY 147 CTGCAGAGGCGGACACCATGATGACCGCCCGCGCGCTCGAGGTCCCGGTGAGGTC 206
    |||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1921 CTGGAAGCGGACCATGATCCCGGTAAACGAGGACTATTACACCGAAGCGCGCTGGTGTG 1980

QY 207 GACGA 211
    | | |
Db 1981 GCCAA 1985

RESULT 13
US-08-881-784-8
; Sequence 8, Application US/08881784
; Patent No. 6083731
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B.
; APPLICANT: Lupien, Shari L.
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
; THE PRODUCTION OF LIMONENE HYDROXYLASES
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
; ADDRESS: 1420 Fifth Avenue, Suite 2800
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,784
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Shelton, Dennis K.
; REGISTRATION NUMBER: 26,997
; REFERENCE/DOCKET NUMBER: WSUR19777
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 224-0718
; TELEFAX: (206) 224-0779
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
```

;
; LENGTH: 1665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; ORIGINAL SOURCE:
; ORGANISM: Mentha x piperita
; IMMEDIATE SOURCE:
; CLONE: ppm17
; US-08-881-784-8

Query Match 18.7%; Score 40; DB 3; Length 1665;
Best Local Similarity 50.5%; Pred. No. 0.23;
Matches 97; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 23 GCGCCCGTGGCGGTTACAAAGCTGTAACAAAGCTGGTCTTGGCGGTGCCAACCCGGC 82
Db 270 GTCCCGGAGGCCACGAAGAGGCGATGAAGCTGTGGACCGCGCTGCGCGGACCGGTT 329
Qy 83 TCTGGTGACTGCCACCGCTCACCGAGGAGACGTGTCGCCACCATCGGTACCTGGT 142
Db 330 CGAGAGCATCGGGAGAGATCATGTGTAGCAACGACATCATCTTCAGCCCCCTA 389
Qy 143 GCGCCTGCACGAGGCCACGACCATGACCGCCCGCGGCGCTCGAGGTCCCGGTGCA 202
Db 390 CAGCGTGACCTGGCGCCAGATGCGGAGATCTGCGTCTCCGAGCTCTCAGCGCCCGCAA 449
Qy 203 GGTCCGACGACAT 214
Db 450 CGTCCGCTCCTT 461

RESULT 14
US-09-292-768-3
; Sequence 3, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; FILE REFERENCE: wsur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Mentha piperita
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1518)
US-09-292-768-3

Query Match 18.7%; Score 40; DB 4; Length 1665;
Best Local Similarity 50.5%; Pred. No. 0.23;
Matches 97; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 23 GCGCCCGTGGCGGTTACAAAGCTGTAACAAAGCTGGTCTTGGCGGTGCCAACCCGGC 82
Db 270 GTCCCGGAGGCCACGAAGAGGCGATGAAGCTGTGGACCGCGCTGCGCGGACCGGTT 329
Qy 83 TCTGGTGACTGCCACCGCTCACCGAGGAGACGTGTCGCCACCATCGGTACCTGGT 142
Db 330 CGAGAGCATCGGGAGAGATCATGTGTAGCAACGACATCATCTTCAGCCCCCTA 389
Qy 143 GCGCCTGCACGAGGCCACGACCATGACCGCCCGCGGCGCTCGAGGTCCCGGTGCA 202
Db 390 CAGCGTGACCTGGCGCCAGATGCGGAGATCTGCGTCTCCGAGCTCTCAGCGCCCGCAA 449

Qy 203 GGTCCGACGACAT 214
Db 450 CGTCCGCTCCTT 461

RESULT 15
US-09-292-768-67
; Sequence 67, Application US/09292768
; Patent No. 6194185
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lupien, Shari L
; APPLICANT: Karp, Frank
; TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR THE PRODUCTION OF
; FILE REFERENCE: wsur13463
; CURRENT APPLICATION NUMBER: US/09/292,768
; EARLIER FILING DATE: 1999-04-14
; EARLIER APPLICATION NUMBER: 08/881,784
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 67
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1665)
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding
; OTHER INFORMATION: limonene-3-hydroxylase variant
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1518)
US-09-292-768-67

Query Match 18.7%; Score 40; DB 4; Length 1665;
Best Local Similarity 50.5%; Pred. No. 0.23;
Matches 97; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 23 GCGCCCGTGGCGGTTACAAAGCTGTAACAAAGCTGGTCTTGGCGGTGCCAACCCGGC 82
Db 270 GTCCCGGAGGCCACGAAGAGGCGATGAAGCTGTGGACCGCGCTGCGCGGACCGGTT 329
Qy 83 TCTGGTGACTGCCACCGCTCACCGAGGAGACGTGTCGCCACCATCGGTACCTGGT 142
Db 330 CGAGAGCATCGGGAGAGATCATGTGTAGCAACGACATCATCTTCAGCCCCCTA 389
Qy 143 GCGCCTGCACGAGGCCACGACCATGACCGCCCGCGGCGCTCGAGGTCCCGGTGCA 202
Db 390 CAGCGTGACCTGGCGCCAGATGCGGAGATCTGCGTCTCCGAGCTCTCAGCGCCCGCAA 449
Qy 203 GGTCCGACGACAT 214
Db 450 CGTCCGCTCCTT 461

Search completed: November 12, 2002, 18:45:22
Job time : 2221.7 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 ; Search time 140.151 Seconds
(without alignments)
3438.621 Million cell updates/sec

Title: US-09-697-123B-7
Perfect score: 214
Sequence: 1 tcaagagagcgctacgac.....cggctcaggctcagacat 214

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	214	100.0	214	22	AAS05207
2	186.8	87.3	214	22	AAS05221
3	153.6	71.8	208	22	AAS05219
4	146.8	68.6	214	22	AAS05214
5	142.4	66.5	208	22	AAS05215
6	140.4	65.6	211	22	AAS05218
7	139.2	65.0	208	22	AAS05201
8	139.2	65.0	208	22	AAS05211
9	137.6	64.3	208	22	AAS05210

10	136	63.6	208	22	AAS05202	Mycobacterium gord
11	134.4	62.8	208	22	AAS05217	Mycobacterium cels
12	132.8	62.1	208	22	AAS05203	Mycobacterium gord
13	132.8	62.1	208	22	AAS05208	Mycobacterium kans
14	128	59.8	208	22	AAS05224	Mycobacterium xeno
15	124.8	58.3	208	22	AAS05205	Mycobacterium tube
16	124.8	58.3	208	22	AAS05206	Mycobacterium terr
17	124.8	58.3	208	22	AAS05216	Mycobacterium bovi
18	124.8	58.3	970	17	AAT09676	Mycobacterium tube
19	124.8	58.3	3519	22	AAH51976	Mycobacterium tube
20	124.8	58.3	3534	22	AAH02079	Mycobacterium tube
21	124.8	58.3	3853	21	AAA74651	Mycobacterium tube
22	124.8	58.3	3853	21	AAA89994	M. tuberculosis rp
23	123.2	57.6	208	22	AAS05222	Mycobacterium affi
24	120.8	56.4	207	22	AAH05212	Mycobacterium szul
25	120	56.1	205	22	AAH05220	Mycobacterium intr
26	118.4	55.3	208	22	AAH05223	Mycobacterium haem
27	116	54.2	207	22	AAH05204	Mycobacterium gord
28	106.2	49.6	223	22	AAH05209	Mycobacterium scro
29	104	48.6	3447	14	AAQ51532	M. leprae rpoB gene
30	99.8	46.6	223	22	AAH05213	Mycobacterium gast
31	88.4	41.3	3495	22	AAH65512	C glutamicum codin
32	88.4	41.3	34980	22	AAH68525	Propionibacterium
33	56.8	26.5	27426	23	AAH59541	Mycobacterium tube
34	43.2	20.2	9210	22	AAH52046	Human ORFX polynuc
35	43.2	20.2	4411529	22	AAI99682	Pseudorabies virus
36	42.2	19.7	318	24	ABN20109	Maize nitrite redu
37	42	19.6	5895	23	ABN87076	groEL-1 gene. Str
38	41.6	19.4	1844	12	AAQ12000	groEL-1 gene. Str
39	41.2	19.3	1620	13	AAQ22482	Thermus sp. FD3041
40	41.2	19.3	2668	13	AAQ22485	Sphingomonas capsu
41	41	19.2	1506	20	AAH76394	S. capsulata ifO12
42	41	19.2	1926	20	AAH82522	Nucleotide sequenc
43	41	19.2	3000	20	AAH84399	Human prostate exp
44	41	19.2	18560	21	AAH58472	
45	40.8	19.1	2076	23	ABV22537	

ALIGNMENTS

RESULT 1
AAS05207
ID AAS05207 standard; DNA; 214 BP.
AC AAS05207;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium chelonae rpoB gene fragment.
DE Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RELP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX Mycobacterium chelonae.
XX WO200131061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
XX WPI; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism

XX PS Claim 1; Page 42; 50pp; English.

XX CC The present sequence for Mycobacterium chelonae rpoB gene

XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from

XX CC various Mycobacterial species. These rpoB gene fragments can be used

XX CC in the diagnosis and identification of Mycobacterium species using a

XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)

XX CC method. The method comprises obtaining a restriction fragment length

XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,

XX CC amplifying and digesting the DNA fragment from the microorganism to

XX CC be identified and comparing the RFLP patterns from the known rpoB gene

XX CC fragments with the unidentified fragment. The rpoB gene fragments

XX CC are useful to identify a wide range of Mycobacterium species, e.g. for

XX CC diagnosis or to obtain epidemiological and pathogenesis information for

XX CC selection of appropriate therapies, including M. tuberculosis, M. leprae

XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected

XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene

XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR

XX CC required), and can differentiate between many species in a single

XX CC experiment, including those difficult to distinguish by usual biochemical

XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for

XX CC detecting specific Mycobacterial species.

XX SQ Sequence 214 BP; 41 A; 72 C; 72 G; 29 T; 0 other;

Query Match 100.0%; Score 214; DB 22; Length 214;

Best Local Similarity 100.0%; Pred. No. 9e-39;

Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGGTACGACCTGGCCCGCGTGGCGGTACAGGTGAACAAGAGCTGG 60

DB 1 TCAAGGAGAAGCGGTACGACCTGGCCCGCGTGGCGGTACAGGTGAACAAGAGCTGG 60

QY 61 GTCTTGGCGGTGCCAACCGGCTCTGTGACTGCCACCGCTCACCGAGAGACGCTCG 120

DB 61 GTCTTGGCGGTGCCAACCGGCTCTGTGACTGCCACCGCTCACCGAGGAGACGCTCG 120

QY 121 TCGCCACCATCGGTACCTGTGTGGCGCTGCACGAGGGCCAGACCATGACCGCCCCCG 180

DB 121 TCGCCACCATCGGTACCTGTGTGGCGCTGCACGAGGGCCAGACCATGACCGCCCCCG 180

QY 181 GCGGCTCGAGTCCCGTCCAGTTCGAGTCCAGCAT 214

DB 181 GCGGCTCGAGTCCCGTCCAGTTCGAGTCCAGCAT 214

RESULT 2

AAS05221

ID AAS05221 standard; DNA; 214 BP.

XX AC AAS05221;

XX DT 07-SEP-2001 (first entry)

XX DE Mycobacterium abscessus rpoB gene fragment.

XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

XX KW PCR-restriction fragment length polymorphism analysis; ds.

XX OS Mycobacterium abscessus.

XX PN WO200131061-A1.

XX PD 03-MAY-2001.

XX PF 27-OCT-2000; 2000WO-KR01223.

XX PR 27-OCT-1999; 99KR-0046795.

XX PA (ERUM-) ERUME BIOTECH CO LTD.

XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX DR WPI; 2001-300520/31.

XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for

XX PT diagnosis and identification of many mycobacterial species by

XX PT restriction fragment length polymorphism

XX PS Claim 1; Page 46; 50pp; English.

XX CC The present sequence for Mycobacterium abscessus rpoB gene

XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from

XX CC various Mycobacterial species. These rpoB gene fragments can be used

XX CC in the diagnosis and identification of Mycobacterium species using a

XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)

XX CC method. The method comprises obtaining a restriction fragment length

XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,

XX CC amplifying and digesting the DNA fragment from the microorganism to

XX CC be identified and comparing the RFLP patterns from the known rpoB gene

XX CC fragments with the unidentified fragment. The rpoB gene fragments

XX CC are useful to identify a wide range of Mycobacterium species, e.g. for

XX CC diagnosis or to obtain epidemiological and pathogenesis information for

XX CC selection of appropriate therapies, including M. tuberculosis, M. leprae

XX CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected

XX CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene

XX CC fragments is rapid, precise, simple and cost effective (only 1 PCR

XX CC required), and can differentiate between many species in a single

XX CC experiment, including those difficult to distinguish by usual biochemical

XX CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for

XX CC detecting specific Mycobacterial species.

XX SQ Sequence 214 BP; 45 A; 69 C; 72 G; 28 T; 0 other;

Query Match 87.3%; Score 186.8; DB 22; Length 214;

Best Local Similarity 92.1%; Pred. No. 9.6e-33;

Matches 197; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGGTACGACCTGGCCCGCGTGGCGGTACAGGTGAACAAGAGCTGG 60

DB 1 TCAAGGAGAAGCGGTACGACCTGGCCCGCGTGGCGGTACAGGTGAACAAGAGCTGG 60

QY 61 GTCTTGGCGGTGCCAACCGGCTCTGTGACTGCCACCGCTCACCGAGGAGACGCTCG 120

DB 61 GCCTGGCGGCACCAATCCGGCTCAGGTGACCAACCACTCACCGAGGAGACGCTCG 120

QY 121 TCGCCACCATCGGTACCTGTGTGGCGCTGCACGAGGGCCAGACCATGACCGCCCCCG 180

DB 121 TCGCCACCATCGGTACCTGTGTGGCGCTGCACGAGGGCCAGACCATGACCGCCCCCG 180

QY 181 GCGGCTCGAGTCCCGTCCAGTTCGAGTCCAGCAT 214

DB 181 GCGGCTCGAGTCCCGTCCAGTTCGAGTTCGAGCAT 214

RESULT 3

AAS05219

ID AAS05219 standard; DNA; 208 BP.

XX AC AAS05219;

XX DT 07-SEP-2001 (first entry)

XX DE Mycobacterium fortuitum rpoB gene fragment.

XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

XX KW PCR-restriction fragment length polymorphism analysis; ds.

XX OS Mycobacterium fortuitum.

XX PN WO200131061-A1.

XX PD 03-MAY-2001.

XX PF 27-OCT-2000; 2000WO-KR01223.

DE Mycobacterium avium rpoB gene fragment.
 XX
 KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
 XX PCR-restriction fragment length polymorphism analysis; ds.
 OS Mycobacterium avium.
 XX
 PN WO200131061-A1.
 XX
 PN 03-MAY-2001.
 XX
 PF 27-OCT-2000; 2000WO-KR01223.
 XX
 PR 27-OCT-1999; 99KR-0046795.
 XX
 PA (ERUM-) ERUME BIOTECH CO LTD.
 XX
 PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
 XX
 DR WPI; 2001-300520/31.
 XX
 XX New DNA fragments from the rpoB gene of mycobacteria, useful for
 PT diagnosis and identification of many mycobacterial species by
 PT restriction fragment length polymorphism -
 XX
 PS Claim 1; Page 44; 50pp; English.
 XX
 CC The present sequence for Mycobacterium avium rpoB gene
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
 CC various Mycobacterium species. These rpoB gene fragments can be used
 CC in the diagnosis and identification of Mycobacterium species using a
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)
 CC method. The method comprises obtaining a restriction fragment length
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
 CC amplifying and digesting the DNA fragment from the microorganism to
 CC be identified and comparing the RFLP patterns from the known rpoB gene
 CC fragments with the unidentified fragment. The rpoB gene fragments
 CC are useful to identify a wide range of Mycobacterium species, e.g. for
 CC diagnosis or to obtain epidemiological and pathogenesis information for
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR
 CC required), and can differentiate between many species in a single
 CC experiment, including those difficult to distinguish by usual biochemical
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
 CC detecting specific Mycobacterium species.
 XX
 SQ Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;
 Query Match 66.5%; Score 142.4; DB 22; Length 208;
 Best Local Similarity 82.7%; Pred. No. 6.6e-23;
 Matches 177; Conservative 0; Mismatches 31; Indels 6; Gaps 1;
 QY 1 TCAAGGAGAGCGCTACGACCTGGCCGGTGGCCGGTACAGGTGAACAAGAGCTGG 60
 DB 1 TCAAGGAGAGCGCTACGACCTGGCCGGTGGCCGGTACAGGTGAACAAGAGCTGG 60
 QY 61 GTCTTGGCGGTGCCAACCCCGGCTCTGGTGACTGCCACACAGCTCACCGAGGAGAGCTCG 120
 DB 61 GCCTGACCGCGGTGAGCC-----GATCACAGCTCGAGCTGACCGAGGAGAGCTCG 114
 QY 121 TCGCCACCATCGGTACCTGGTGGCGCTGCACGAGGGCCAGACCATGATGACCGCCCGC 180
 DB 115 TCGCCACCATCGAGTACCTGGTGGCGCTGCACGAGGGTGCAGCCACGATGACCGCTCCCG 174
 QY 181 CGGCCCTCGAGGTCCCGGTGCGAGCTGCAGACAT 214
 DB 175 CGGGCATCGAGGTCCCGGTGGAGACCGACGACAT 208
 RESULT 6
 AAS05218

AA050218 standard; DNA; 211 BP.
 AAS05218;
 07-SEP-2001 (first entry)
 Mycobacterium flavescens rpoB gene fragment.
 Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
 PCR-restriction fragment length polymorphism analysis; ds.
 Mycobacterium flavescens.
 WO200131061-A1.
 03-MAY-2001.
 27-OCT-2000; 2000WO-KR01223.
 27-OCT-1999; 99KR-0046795.
 (ERUM-) ERUME BIOTECH CO LTD.
 Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
 WPI; 2001-300520/31.
 New DNA fragments from the rpoB gene of mycobacteria, useful for
 diagnosis and identification of many mycobacterial species by
 restriction fragment length polymorphism -
 Claim 1; Page 45; 50pp; English.
 The present sequence for Mycobacterium flavescens rpoB gene
 fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
 various Mycobacterium species. These rpoB gene fragments can be used
 in the diagnosis and identification of Mycobacterium species using a
 novel PCR-restriction fragment length polymorphism analysis (PRA)
 method. The method comprises obtaining a restriction fragment length
 polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
 amplifying and digesting the DNA fragment from the microorganism to
 be identified and comparing the RFLP patterns from the known rpoB gene
 fragments with the unidentified fragment. The rpoB gene fragments
 are useful to identify a wide range of Mycobacterium species, e.g. for
 diagnosis or to obtain epidemiological and pathogenesis information for
 selection of appropriate therapies, including M. tuberculosis, M. leprae
 and non-tuberculous mycobacteria (NTM) encountered in subjects infected
 with human immunodeficiency virus (HIV). Analysis of the rpoB gene
 fragments is rapid, precise, simple and cost effective (only 1 PCR
 required), and can differentiate between many species in a single
 experiment, including those difficult to distinguish by usual biochemical
 tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
 detecting specific Mycobacterium species.
 Sequence 211 BP; 49 A; 65 C; 70 G; 27 T; 0 other;
 Query Match 65.6%; Score 140.4; DB 22; Length 211;
 Best Local Similarity 81.8%; Pred. No. 1.8e-22;
 Matches 175; Conservative 0; Mismatches 36; Indels 3; Gaps 1;
 QY 1 TCAAGGAGAGCGCTACGACCTGGCCGGTGGCCGGTACAGGTGAACAAGAGCTGG 60
 DB 1 TCAAGGAGAGCGCTACGACCTGGCCGGTGGCCGGTACAGGTGAACAAGAGCTGG 60
 QY 61 GTCTTGGCGGTGCCAACCCCGGCTCTGGTGACTGCCACACAGCTCACCGAGGAGAGCTCG 120
 DB 61 GCAT---CACCGAGAAACCCGGCCGACACACCTCGACACCGTGCACGAGGAGCTCG 117
 QY 121 TCGCCACCATCGGTACCTGGTGGCGCTGCACGAGGGCCAGACCATGATGACCGCCCGC 180
 DB 118 TCGCCACCATCGAGTACCTGGTGGCGCTGCATCAGGGCGCAAGACGATGACCGCTCCCG 177
 QY 181 CGGGCTCGAGGTTCCTCGGTCGAGGTTCGACGACAT 214

QY	121	TCGCCACCATCGGGTACCTGCTGGCGCTGTCACGAGGCCAGACCCACGATGACGCCGCCCG	180
DB	115	TCGCCACCATCGAGTACCTGCTGGCGCTGTCACGAGGCCAGACCCACGATGACGCCGCCCG	174
QY	181	GGGGCTCGAGTCCCGTCCGAGTCCGAGTCCGACGACAT	214
DB	175	GGGGACCGAGTCCCGTTCGAGACCGACGACAT	208
RESULT 8			
ID	AA05211 standard; DNA; 208 BP.		
XX	AA05211;		
DT	07-SEP-2001 (first entry)		
XX	Mycobacterium marinum rpoB gene fragment.		
XX	Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP		
KW	PCR-restriction fragment length polymorphism analysis; ds.		
XX	Mycobacterium marinum.		
OS	WO200131061-A1.		
PN	XX		
XX	03-MAY-2001.		
XX	27-OCT-2000; 2000WO-KR01223.		
PF	XX		
XX	27-OCT-1999; 99KR-0046795.		
PR	XX		
XX	(BRUM-) ERUME BIOTECH CO LTD.		
PA	XX		
XX	Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;		
FI	XX		
XX	WPI; 2001-300520/31.		
DR	XX		
XX	New DNA fragments from the rpoB gene of mycobacteria, useful for		
PT	diagnosis and identification of many mycobacterial species by		
PT	restriction fragment length polymorphism		
XX	Claim 1; Page 43; 50pp; English.		
XX	The present sequence for Mycobacterium marinum rpoB gene		
XX	fragment is 1 of 24 rpoB gene fragments (AA05201-AA05224) from		
CC	various Mycobacterial species. These rpoB gene fragments can be used		
CC	in the diagnosis and identification of Mycobacterium species using a		
CC	novel PCR-restriction fragment length polymorphism analysis (PRA)		
CC	method. The method comprises obtaining a restriction fragment length		
CC	polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,		
CC	amplifying and digesting the DNA fragment from the microorganism to		
CC	be identified and comparing the RFLP patterns from the known rpoB gene		
CC	fragments with the unidentified fragment. The rpoB gene fragments		
CC	are useful to identify a wide range of Mycobacterium species, e.g. for		
CC	diagnosis or to obtain epidemiological and pathogenesis information for		
CC	selection of appropriate therapies, including M. tuberculosis, M. leprae		
CC	and non-tuberculous mycobacteria (NTM) encountered in subjects infected		
CC	with human immunodeficiency virus (HIV). Analysis of the rpoB gene		
CC	fragments is rapid, precise, simple and cost effective (only 1 PCR		
CC	required), and can differentiate between many effective (only 1 PCR		
CC	experiment, including those difficult to distinguish by a single		
CC	tests. Also described are oligonucleotide probes (AA05227-AA05242) for		
CC	detecting specific Mycobacterial species.		
XX	Sequence 208 BP; 45 A; 69 C; 68 G; 26 T; 0 other;		
SQ	XX		
Query Match 65.0%; Score 139.2; DB 22; Length 208;			
Best Local Similarity 81.8%; Pred. No. 3.4e-22;			
Matches 175; Conservative 0; Mismatches 33; Indels 6; Gaps			
QY	1 TCAGGAGAACGCTACGACTGGCGCCGGTGGCGCCGTACAGGTGAACAAGAGCTGG 60		

CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterium species.

XX SQ Sequence 208 BP; 45 A; 62 C; 73 G; 28 T; 0 other;
Query Match 63.6%; Score 136; DB 22; Length 208;
Best Local Similarity 80.8%; Pred. No. 1.7e-21;
Matches 173; Conservative 0; Mismatches 35; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCGCGTGGCGGTGACAAAGTGACAAAGAGCTGG 60
DB 1 TCAAGGAGAGCGCTACGACCTGGCGCGTGGCGGTGACAAAGTGACAAAGAGCTGG 60
QY 61 GTCTTGGCGGTGCAACCGCGCTCTGGTACTGCCACACCGTACCCAGGAGAGCTGG 120
DB 61 GTCTTGGCGGTGCAACCGCGCTCTGGTACTGCCACACCGTACCCAGGAGAGCTGG 114
QY 121 TCGCCACCATCGGCTACCTGGTGGCGCTGCAGAGGGCCAGACACGATGACCGCCCGG 180
DB 115 TAGCCACCATCGAGTACCTGGTGGCGCTGCAGAGGGTCACTGCGCGATGACGGTTCCCG 174
QY 181 GCGGCTCGAGTCCCGTGGAGTGGAGGACAT 214
DB 175 GCGGCGCGAGGTGGCGGTGGAGACCGAGCAT 208

RESULT 11
AAS05217
ID AAS05217 standard; DNA; 208 BP.
XX AC AAS05217;
XX DT 07-SEP-2001 (first entry)
XX DE Mycobacterium celatum rpoB gene fragment.
XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX KW PCR-restriction fragment length polymorphism analysis; ds.
XX OS Mycobacterium celatum.
XX PN WO200131061-A1.
XX PD 03-MAY-2001.
XX PF 27-OCT-2000; 2000WO-KR01223.
XX PR 27-OCT-1999; 99KR-0046795.
XX PA (ERUM-) ERUME BIOTECH CO LTD.
XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX DR WPI; 2001-300520/31.
XX PS New DNA fragments from the rpoB gene of mycobacteria, useful for
XX PT diagnosis and identification of many mycobacterial species by
XX PT restriction fragment length polymorphism -
XX PS Claim 1; Page 45; 50pp; English.

CC The present sequence for Mycobacterium celatum rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterium species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments

CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterium species.

XX SQ Sequence 208 BP; 48 A; 71 C; 64 G; 25 T; 0 other;
Query Match 62.8%; Score 134.4; DB 22; Length 208;
Best Local Similarity 80.4%; Pred. No. 3.9e-21;
Matches 172; Conservative 0; Mismatches 36; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCGCGTGGCGGTGACAAAGTGACAAAGAGCTGG 60
DB 1 TCAAGGAGAGCGCTACGACCTGGCGCGTGGCGGTGACAAAGTGACAAAGAGCTGG 60
QY 61 GTCTTGGCGGTGCCAACCGCGCTCTGGTACTGCCACACCGTACCCAGGAGAGCTGG 120
DB 61 GCCT-----GAACACCGCGTCCCGATCACGACGACCTCTGACCGAGAGAGCTGG 114
QY 121 TCGCCACCATCGGCTACCTGGTGGCGCTGCAGAGGGCCAGACGATGACCGCCCGG 180
DB 115 TCGCCACCATCGAGTACCTGGTGGCGCTGCAGAGGGCCAGACGATGACCGCCCGG 174
QY 181 GCGGCTCGAGTCCCGTGGAGTGGAGGACAT 214
DB 175 GCGGAGTTCGAGTGGCGGTGGAAACCGAGCAT 208

RESULT 12
AAS05203
ID AAS05203 standard; DNA; 208 BP.
XX AC AAS05203;
XX DT 07-SEP-2001 (first entry)
XX DE Mycobacterium gordonae type III rpoB gene fragment.
XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX KW PCR-restriction fragment length polymorphism analysis; ds.
XX OS Mycobacterium gordonae type III.
XX PN WO200131061-A1.
XX PD 03-MAY-2001.
XX PF 27-OCT-2000; 2000WO-KR01223.
XX PR 27-OCT-1999; 99KR-0046795.
XX PA (ERUM-) ERUME BIOTECH CO LTD.
XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX DR WPI; 2001-300520/31.
XX PS New DNA fragments from the rpoB gene of mycobacteria, useful for
XX PT diagnosis and identification of many mycobacterial species by
XX PT restriction fragment length polymorphism -
XX PS Claim 1; Page 41; 50pp; English.

CC The present sequence for Mycobacterium gordonae type III rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterium species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a

CC novel PCR-restriction fragment length polymorphism analysis (PRA)
 CC method. The method comprises obtaining a restriction fragment length
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
 CC amplifying and digesting the DNA fragment from the microorganism to
 CC be identified and comparing the RFLP patterns from the known rpoB gene
 CC fragments with the unidentified fragment. The rpoB gene fragments
 CC are useful to identify a wide range of Mycobacterium species, e.g. for
 CC diagnosis or to obtain epidemiological and pathogenesis information for
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR
 CC required), and can differentiate between many species in a single
 CC experiment, including those difficult to distinguish by usual biochemical
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
 CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 46 A; 69 C; 63 G; 30 T; 0 other;

Query Match 62.1%; Score 132.8; DB 22; Length 208;
 Best Local Similarity 79.9%; Pred. No. 8.9e-21;
 Matches 171; Conservative 0; Mismatches 37; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGGTACGACCTGGCCCGGTGGCCGGTACAAGGTCAACAAGAGCTGG 60
 DB 1 TCAAGGAGAGCGGTACGACCTGGCCCGGTGGCCGGTACAAGGTCAACAAGAGCTGG 60
 QY 61 GTCTTGGCGGTGCAACCCGGCTCTGTGACTGCCACCGCTCACCGAGGAGAGCTGG 120
 DB 61 GCCTGCGAGTCCGGC-----GATCCGATCACCGCTCCAGCTGACCGAAGAGAGCTGG 114
 QY 121 TCGCCACCATCGGTACCTGGTGGCTGCACGAGGGCCAGACGATGACCGCCCGG 180
 DB 115 TCGCCACCATCGGTACCTGGTGGCTGCACGAGGGTTCAGCAGATGACCGTTCGG 174
 QY 181 CGCGCCTCGAGGTCCCGGTGAGGTGCGACGACAT 214
 DB 175 CGCGCAGCGAGGTTCGGGTGGAGACGACGACAT 208

RESULT 13

AAS05208
 ID AAS05208 standard; DNA; 208 BP.

AC AAS05208;

DT 07-SEP-2001 (first entry)

DE Mycobacterium kansasii rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
 KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium kansasii.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
 PI WPI; 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for
 PT diagnosis and identification of many mycobacterial species by
 PT restriction fragment length polymorphism

XX

PS Claim 1; Page 42; 50pp; English.

XX The present sequence for Mycobacterium kansasii rpoB gene
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05242) from
 CC various Mycobacterial species. These rpoB gene fragments can be used
 CC in the diagnosis and identification of Mycobacterium species using a
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)
 CC method. The method comprises obtaining a restriction fragment length
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
 CC amplifying and digesting the DNA fragment from the microorganism to
 CC be identified and comparing the RFLP patterns from the known rpoB gene
 CC fragments with the unidentified fragment. The rpoB gene fragments
 CC are useful to identify a wide range of Mycobacterium species, e.g. for
 CC diagnosis or to obtain epidemiological and pathogenesis information for
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR
 CC required), and can differentiate between many species in a single
 CC experiment, including those difficult to distinguish by usual biochemical
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
 CC detecting specific Mycobacterial species.

XX Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;

Query Match 62.1%; Score 132.8; DB 22; Length 208;
 Best Local Similarity 79.9%; Pred. No. 8.9e-21;
 Matches 171; Conservative 0; Mismatches 37; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGGTACGACCTGGCCCGGTGGCCGGTACAAGGTCAACAAGAGCTGG 60
 DB 1 TCAAGGAGAGCGGTACGACCTGGCCCGGTGGCCGGTACAAGGTCAACAAGAGCTGG 60
 QY 61 GTCTTGGCGGTGCAACCCGGCTCTGTGACTGCCACCGCTCACCGAGGAGAGCTGG 120
 DB 61 GCCTGCGAGTCCGGC-----GAACACCAATCATCCGATCACCGAGGAGAGCTGG 114
 QY 121 TCGCCACCATCGGTACCTGGTGGCTGCACGAGGGCCAGACGATGACCGCCCGG 180
 DB 115 TCGCCACCATCGGTACCTGGTGGCTGCACGAGGGTTCAGCAGATGACCGTTCGG 174
 QY 181 CGCGCCTCGAGGTCCCGGTGAGGTGCGACGACAT 214
 DB 175 CGCGGTCGAGGTGCGCGGTGGAAACCGACGACAT 208

RESULT 14

AAS05224
 ID AAS05224 standard; DNA; 208 BP.

AC AAS05224;

DT 07-SEP-2001 (first entry)

DE Mycobacterium xenopi rpoB gene fragment.

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
 KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium xenopi.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
 PI

DR	WPI; 2001-300520/31.
XX	
PT	New DNA fragments from the rpoB gene of mycobacteria, useful for
PT	diagnosis and identification of many mycobacterial species by
PT	restriction fragment length polymorphism -
XX	
PS	Claim 1; Page 47; 50pp; English
XX	
CC	The present sequence for Mycobacterium xenopi rpoB gene
CC	fragment is 1 of 24 rpoB gene fragments (AAS05201-AA05224) from
CC	various Mycobacterial species. These rpoB gene fragments can be used
CC	in the diagnosis and identification of Mycobacterium species using a
CC	novel PCR-restriction fragment length polymorphism analysis (PRA)
CC	method. The method comprises obtaining a restriction fragment length
CC	polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC	amplifying and digesting the DNA fragment from the microorganism to
CC	be identified and comparing the RFLP patterns from the known rpoB gene
CC	fragments with the unidentified fragment. The rpoB gene fragments
CC	are useful to identify a wide range of Mycobacterium species, e.g. for
CC	diagnosis or to obtain epidemiological and pathogenesis information for
CC	selection of appropriate therapies, including M. tuberculosis, M. leprae
CC	and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC	with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC	fragments is rapid, precise, simple and cost effective (only 1 PCR
CC	required), and can differentiate between many species in a single
CC	experiment), including those difficult to distinguish by usual biochemical
CC	tests. Also described are oligonucleotide probes (AAS05227-AA05242) for
CC	detecting specific Mycobacterial species.
XX	
SQ	Sequence 208 BP; 51 A; 65 C; 67 G; 25 T; 0 other;
	Query Match 59.8%; Score 128; DB 22; Length 208;
	Best Local Similarity 78.5%; Pred. No. 1e-19;
	Matches 168; Conservative 0; Mismatches 40; Indels 6; Gaps 1
Qy	1 TCAAGGAGAAGCGCTACGACCTGGCCGGTGGCCGGGTACAAGGTGAACAAGAAGCTGG 60
Db	1 TCAAGGAGAAGCGCTACGACCTGGCCGGTGGCCGGGTACAAGGTCAACAGAAACTCG 60
Qy	61 GTCTTGCGGTGCCAACCCGGCTCTGGTAGCTGCACCACCGCTCACCGAGAAGAGCTG 120
Db	61 GGCTGAACACCGAGAATGCGCC-----AACCCACGACGCTCAGCGAAGAGGACGTCG 114
Qy	121 TCGCCACCATTCGGGTACCTGGTGGCGCTGCAGAGGGCCAGCACCATGACCGCCCCCG 180
Db	115 TCGCCACCATTCGAATACCTGGTGGCTTGCAGAGGGGCACGCCCATGAAGTCCCCG 174
Qy	181 CGCGCTCGAGTCCCGGTTCGAGTTCGACGACAT 214
Db	175 GTGGCGTTCGAGTGGCGGTGGAGACCGACGACAT 208
RESULT 15	
AAS05205	
ID	AAS05205 standard; DNA; 208 BP.
XX	AAS05205;
XX	
DT	07-SEP-2001 (first entry)
XX	
XX	Mycobacterium tuberculosis rpoB gene fragment.
DE	
XX	
KW	Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW	PCR-restriction fragment length polymorphism analysis; ds.
OS	
XX	Mycobacterium tuberculosis.
PX	
NN	WO200131061-A1.
XX	
PD	03-MAY-2001.
PF	
XX	27-OCT-2000; 2000WO-KR01223.
XX	

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 : Search time 654.37 Seconds
(without alignments)
9517.553 Million cell updates/sec

Title: US-09-697-123b-7

Perfect score: 214

Sequence: 1 tcaaggagaagcgctacgac.....ccggtcgaggtcgacgacat 214

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rtd.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	132.8	62.1	3752	1	MSU24494	U24494 Mycobacteri
2	124.8	58.3	610	1	MTU318818	AJ318818 Mycobacte
3	124.8	58.3	610	1	MTU318819	AJ318819 Mycobacte
4	124.8	58.3	616	1	MTU318813	AJ318813 Mycobacte
5	124.8	58.3	618	1	MTU318815	AJ318815 Mycobacte
6	124.8	58.3	618	1	MTU318817	AJ318817 Mycobacte
7	124.8	58.3	633	1	MTU318814	AJ318814 Mycobacte
8	124.8	58.3	637	1	MTU318816	AJ318816 Mycobacte
9	124.8	58.3	639	1	MTU318821	AJ318821 Mycobacte
10	124.8	58.3	970	6	150706	150706 Sequence 1
11	124.8	58.3	3534	6	AX111339	AX111339 Sequence
12	124.8	58.3	3853	1	MTU12205	U12205 Mycobacteri
13	124.8	58.3	5084	1	MSGRPO8	L27989 Mycobacteri
14	124.8	58.3	19352	1	AE006964	AE006964 Mycobacte
15	124.8	58.3	19770	1	MTU1376	295972 Mycobacteri
16	104	48.6	3447	6	AR067447	AR067447 Sequence
17	104	48.6	37617	1	MLB1790G	Z14314 M.leprae ge
18	104	48.6	348950	1	MLEPRTN7	AL583923 Mycobacte
19	103.2	48.2	3941	1	AF242549	AF242549 Mycolato
20	88.4	41.3	3495	6	AX120631	AX120631 Sequence
21	88.4	41.3	328050	1	AP005375	AP005375 Corynebac
22	88.4	41.3	349980	6	AX127144	AX127144 Sequence
23	84	39.3	32923	1	SCD82	AL160431 Streptomy
24	47	22.0	11103	1	AE001944	AE001944 Deinococc
25	46.8	21.9	7854	1	AX024345	AX024345 Sequence
26	46.8	21.9	7854	6	AX024238	AX024238 Sequence
27	46.8	21.9	47713	1	AX024320	AX024320 Sequence
28	46.8	21.9	47713	6	AX024213	AX024213 Sequence
29	46	21.5	3350	1	SCAPTRELA	X87267 S.coelicolo
30	44.8	20.9	101365	2	AF004748	AF004748 Oryza sat
31	44.8	20.9	152263	2	AF005412	AF005412 Oryza sat
32	44.6	20.8	329100	1	SME59187	AL591787 S.inorhizo
33	44.4	20.7	4068	1	SCPPGPP	X92520 S.coelicolo
34	44.4	20.7	38640	1	SCL2	AL137778 Streptomy
35	44	20.6	30561	1	SC1F2	AL031350 Streptomy
36	44	20.6	34023	1	SC2G2	AL445963 Streptomy
37	43.6	20.4	1810	1	AF222752	AF222752 Bradyrhiz
38	43.6	20.4	12277	1	AF340166	AF340166 Streptomy
39	43.6	20.4	37159	2	AC100259	AC100259 Mus muscu
40	43.6	20.4	134188	2	AC120983	AC120983 Oryza sat
41	43.4	20.3	4746	1	ABHISHAFE	X61207 A.brasilens
42	43.2	20.2	17849	1	AE007096	AE007096 Mycobacte
43	43.2	20.2	33818	1	MTCY159	Z83863 Mycobacteri
44	43	20.1	2904	1	SREFRS	Y14317 Streptomyce
45	43	20.1	118733	2	AP003885	AP003885 Oryza sat

ALIGNMENTS

RESULT 1
MSU24494
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

MSU24494
Mycobacterium smegmatis DNA polymerase (rpoB) gene, complete cds.
U24494
U24494.1 GI:790347

Mycobacterium smegmatis.
Mycobacterium smegmatis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
1 (bases 1 to 3752)

REFERENCE
AUTHORS
TITLE
Hetherington,S.V., Watson,A.S. and Patrick,C.C.
Sequence and analysis of the rpoB gene of Mycobacterium smegmatis

JOURNAL Antimicrob. Agents Chemother. 39 (9), 2164-2166 (1995)
MEDLINE 96050766
PUBMED 8540740
REFERENCE 2 (bases 1 to 3752)
AUTHORS Hetherington,S.V.
TITLE Direct Submission
JOURNAL Submitted (11-APR-1995) Seth V. Hetherington, Infectious Diseases,
St. Jude Children's Research Hospital, 332 N. Lauderdale, Memphis,
TN 38101, USA

FEATURES

Location/Qualifiers

1..3752
/organism="Mycobacterium smegmatis"

/db_xref="taxon:1772"

194..3703

/gene="rpoB"

/gene="rpoB"

/codon_start=1

/evidence=experimental

/transl_table=11

/product="DNA polymerase"

/protein_id="AAA91426.2"

/db_xref="gi:7144498"

LDVQDSFMDVGGRRVROAIDRGEENPVGGLAEVLAELSPTEDFSGMSLSFSDPR
FDEVKASDFECKDKDMTYAAPLFVTAEFINNTEGEIKSOTFVMDPFPMTKGTFFIN
GTERVVVSOLVRSFVYFDETDKSTKTLHSVKVIFPGANLEFVDKRDVTGVRID
PKRKQPVTLKALGNLQVFESEIMMGTLKEDTSGDEALLDIYLRKLRPE
PPTKESQATLLLENFFKRYDLARVGVKYNKGLNAGKPIITSTLTEDVYATIE
YLRLHCEQTSMTVGGVFAVESHDIDHFNRLRTVGLIQNIRVLSRMERVVVR
ERMVTDQVEAITPOTLINIRPVAAIKKEFTGTSOLFQMDQNNPLSLGTHKRLSALG
PGGLSRAGLEVLRDHPHSHVRCMPETETPEGNLIGLSLVYARVNPFGFETPYR
KVVSRRVTDQDYLTADEEDPHVVAQANSTPDENGRTEDRVVRKKGVEVFVSADQ
VYMDVSPRQHVSVATAMPIFLEHDDANRALMGANQRAVPLVRSAPLVGTGMELR
AAIDAVTWSHKTQYIEVSADYITVMADDTGTSYLRKRPANSTCANQRPVIDA
RGQAGVATDCTONGEMALKNLLVAMPWEGHNYEDAIIISRLVEEDVLTISI
HIEHEIDARDKILGABEITRIVPKLSDLEADLDEPRIVRGAESREDDELPVGVN
KGETELPDEARLRAIFEGAREVRDTSKLVPHGESKVGIVRFSREDDELPVGVN
VLNVYVAOKRISDQKLAHNGKVGIGKILPVEDMPFLPDGTVDVILNTHGCRV
ELVNGOILETHLHGKAGWNIDVLAGVPPDWSKLPRELYSAPADSTVATPVFDGAQE
GELAGLLGSTLPNRDGEVNVNADKATLFDGRSSGEPFPYPTVGYMILKHLHVD
KIHARSTGPSYMTQOPLGKGAQFGGQRFGEKMECMAQYAAATYQLLELTIKSDDTV
GRVKVYEAIVKGENIPEPGIPESFKVLLKELQSLCNLNEVLSSDGRAIENRDRGDDDL
ERAAANGINLSRNESASVEDLA"

BASE COUNT 706 a 1225 c 1210 g 611 t

Query Match 62.1%; Score 132.8; DB 1; Length 3752;
Best Local Similarity 79.9%; Pred. No. 1.7e-13;
Matches 171; Conservative 0; Mismatches 37; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCGCGGTGGCGGTACAGGTGACCAAGAGCTGG 60
|||||
DB 1068 TCAAGGAGAGCGCTACGACCTGGCGCGGTGGCGGTACAGGTGACCAAGAGCTGG 1127
QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGTGACTGCCACCCAGCTCACCGAGGAAGAGCTGG 120
|||
DB 1128 GCCTCAACCGGGCAAGCC-----GATCACCAGCTCGACCTGACCGAGAGAGCTGG 1181
QY 121 TCGCCACCATCGGTACTGTGGCGCTTCGACGAGGGCCAGACCATGATACCCGCCCG 180
|||
DB 1182 TCGCAGCATCGATACCTGTGGTCTGACGAGGGTTCAGACCTCGATACCCGCTCCCG 1241
QY 181 GCGGCTCGAGTCCCGCTCGAGTTCGACGACAT 214
|||
DB 1242 GTGGCGTCTGAGTTCGCCGCTCGAGTTCGACGACAT 1275

RESULT 2
LOCUS MTU318818 610 bp DNA circular BCT 09-AUG-2002
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1415-97.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AJ318818
AJ318818.1 GI:22208412
RNA polymerase beta subunit; rpoB gene.
Mycobacterium tuberculosis.
Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE
AUTHORS
TITLE

Herrera,L., Jimenez,M.S. and Saez,J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review

JOURNAL
REFERENCE
AUTHORS
TITLE

2 (bases 1 to 610)
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN

FEATURES
source

1..610
Location/Qualifiers
/organism="Mycobacterium tuberculosis"
/isolate="1415-97"
/db_xref="taxon:1773"
1..610
/gene="rpoB"
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87035.1"
/db_xref="GI:22208413"

/translation="LDIYKLRPGPPTKESQATLLNLFKRYDLARVGRYKVNK
KLGLHVGEPITSTLTEDVYATIEYLRLHCEQTSMTVGGVFAVESHDIDHFNRLRTVGLIQNIRVLSRMERVVVR
ERMVTDQVEAITPOTLINIRPVAAIKKEFTGTSOLFQMDQNNPLSLGTHKRLSALG
PGGLSRAGLEVLRDHPHSHVRCMPETETPEGNLIGLSLVYARVNPFGFETPYR
KVVSRRVTDQDYLTADEEDPHVVAQANSTPDENGRTEDRVVRKKGVEVFVSADQ
VYMDVSPRQHVSVATAMPIFLEHDDANRALMGANQRAVPLVRSAPLVGTGMELR
AAIDAVTWSHKTQYIEVSADYITVMADDTGTSYLRKRPANSTCANQRPVIDA
RGQAGVATDCTONGEMALKNLLVAMPWEGHNYEDAIIISRLVEEDVLTISI
HIEHEIDARDKILGABEITRIVPKLSDLEADLDEPRIVRGAESREDDELPVGVN
KGETELPDEARLRAIFEGAREVRDTSKLVPHGESKVGIVRFSREDDELPVGVN
VLNVYVAOKRISDQKLAHNGKVGIGKILPVEDMPFLPDGTVDVILNTHGCRV
ELVNGOILETHLHGKAGWNIDVLAGVPPDWSKLPRELYSAPADSTVATPVFDGAQE
GELAGLLGSTLPNRDGEVNVNADKATLFDGRSSGEPFPYPTVGYMILKHLHVD
KIHARSTGPSYMTQOPLGKGAQFGGQRFGEKMECMAQYAAATYQLLELTIKSDDTV
GRVKVYEAIVKGENIPEPGIPESFKVLLKELQSLCNLNEVLSSDGRAIENRDRGDDDL
ERAAANGINLSRNESASVEDLA"

BASE COUNT
ORIGIN

122 a 191 c 202 g 95 t
Query Match 58.3%; Score 124.8; DB 1; Length 610;
Best Local Similarity 77.6%; Pred. No. 4.9e-12;
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCGCGGTGGCGGTGACAGGTGACCAAGAGCTGG 60
|||||
DB 80 TCAAGGAGAGCGCTACGACCTGGCGCGGTGGCGGTGACAGGTGACCAAGAGCTGG 139
QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGTGACTGCCACCCAGCTCACCGAGGAAGAGCTGG 120
|||
DB 140 GGCTGATGTCGCGGAGCC-----CATCAGCTCGTGCAGCTGACCGAGGAAGAGCTGG 193
QY 121 TCGCCACCATCGGTACTGTGGCGCTTCGACGAGGGCCAGACCATGATACCCGCCCG 180
|||||
DB 194 TGGCCACCATCGAATATCTGTCGCTTCGACGAGGGTTCAGACCATGATACCCGTC 253
QY 181 GCGGCTCGAGTCCCGCTCGAGTTCGACGACAT 214
|||||
DB 254 GCGGCTCGAGTTCGCGGTGGAAACCGACGACAT 287

RESULT 3
LOCUS
DEFINITION

MTU318819 610 bp DNA circular BCT 09-AUG-2002
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1417-97.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AJ318819
AJ318819.1 GI:22208414
RNA polymerase beta subunit; rpoB gene.
Mycobacterium tuberculosis.
Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.

REFERENCE 1
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis isolated in Spain (1996-2001). Description of new alleles into rpoB gene and review

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 610)

AUTHORS Herrera, L.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5, Majadahonda. Madrid. 28220, SPAIN

FEATURES Location/Qualifiers

source 1..610
/organism="Mycobacterium tuberculosis"
/isolate="1417-97"
/db_xref="taxon:1773"

gene 1..610

CDS /gene="rpoB"

<1..>610

/gene="rpoB"

/codon_start=1

/transl_table=11

/product="RNA polymerase beta subunit"

/protein_id="CAC87036.1"

/db_xref="GI:2208415"

/translation="LDIYRKLRPGEPTKESAOITLLENLFFKEKRYDLARVRYKVNK
KLGHVGEPTSTLTEDVATIEYLVRLHEGQTTMTVPGVEVPVETDDIDHFGNR
RLRTVGEIQNIQIRVGSRMERVRMTTQDVEAITPOTLINIRPVVAAIKEFFGTS
QPSQFMQNNPLSGLTHKRLSALGPGLSRERAGLEVRDHP"

BASE COUNT 122 a 191 c 202 g 95 t

ORIGIN

Query Match 58.3%; Score 124.8; DB 1; Length 610;

Best Local Similarity 77.6%; Pred. No. 4.9e-12;

Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGTGGCGCGGTACAAAGGTGAACAAGACTGG 60

Db 80 TCAAGGAGAGCGCTACGACCTGGCCGCGTGGCGCGGTACAAAGGTGAACAAGACTGG 139

QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGGTACTCCACACAGCTCACCGAAGAGACTCG 120

Db 140 GGCTGCATGTCGGCGAGCC-----CATCACGTCGTCGACGCTGACCGAAGAGACTCG 193

QY 121 TCGCCACCATCGGGTACCTGGTGGCGCTGCACGAGGCGCAGACCATGACCGCCCGG 180

Db 194 TGGCCACCATCGAATATCTGGTCCGCTTGACGAGGCTGACCGAAGAGACTCG 253

QY 181 GCGGCTCGAGGTCCCGGTGCGAGGTCGACGACAT 214

Db 254 GCGGCTCGAGGTCCCGGTGCGAACCAGCAGACAT 287

RESULT 4

MTU318813

LOCUS

DEFINITION

616 bp DNA circular BCT 09-AUG-2002

Mycobacterium tuberculosis partial rpoB gene for RNA polymerase

beta subunit, isolate 1763-97.

ACCESSION

AJ318813

VERSION

RNA polymerase beta subunit; rpoB gene.

KEYWORDS

Mycobacterium tuberculosis.

SOURCE

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium

tuberculosis complex.

REFERENCE 1

Herrera, L., Jimenez, M.S. and Saez, J.A.

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis

isolated in Spain (1996-2001). Description of new alleles into rpoB

gene and review

Unpublished

REFERENCE 2 (bases 1 to 610)

Herrera, L.

Direct Submission

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Unpublished

REFERENCE 2 (bases 1 to 616)

Herrera, L.

Direct Submission

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro

Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,

Majadahonda. Madrid. 28220, SPAIN

FEATURES Location/Qualifiers

source 1..616

/organism="Mycobacterium tuberculosis"

/isolate="1763-97"

/db_xref="taxon:1773"

gene 1..615

CDS /gene="rpoB"

<1..>615

/gene="rpoB"

/codon_start=1

/transl_table=11

/product="RNA polymerase beta subunit"

/protein_id="CAC87030.1"

/db_xref="GI:2208403"

/translation="LDIYRKLRPGEPTKESAOITLLENLFFKEKRYDLARVRYKVNK

KLGHVGEPTSTLTEDVATIEYLVRLHEGQTTMTVPGVEVPVETDDIDHFGNR

RLRTVGEIQNIQIRVGSRMERVRMTTQDVEAITPOTLINIRPVVAAIKEFFGTS

QPSQFIFMQNNPLSGLTHKRLSALGPGLSRERAGLEVRDHP"

BASE COUNT 125 a 191 c 201 g 99 t

ORIGIN

Query Match 58.3%; Score 124.8; DB 1; Length 616;

Best Local Similarity 77.6%; Pred. No. 4.9e-12;

Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGTGGCGCGGTACAAAGGTGAACAAGACTGG 60

Db 80 TCAAGGAGAGCGCTACGACCTGGCCGCGTGGCGCGGTACAAAGGTGAACAAGACTGG 139

QY 61 GTCTTGGCGGTGCCAACCCGGCTCTGGTACTCCACACAGCTCACCGAAGAGACTCG 120

Db 140 GGCTGCATGTCGGCGAGCC-----CATCACGTCGTCGACGCTGACCGAAGAGACTCG 193

QY 121 TCGCCACCATCGGGTACCTGGTGGCGCTGCACGAGGCGCAGACCATGACCGCCCGG 180

Db 194 TGGCCACCATCGAATATCTGGTCCGCTTGACGAGGCTGACCGAAGAGACTCG 253

QY 181 GCGGCTCGAGGTCCCGGTGCGAGGTCGACGACAT 214

Db 254 GCGGCTCGAGGTCCCGGTGCGAACCAGCAGACAT 287

RESULT 5

MTU318815

LOCUS

DEFINITION

618 bp DNA circular BCT 09-AUG-2002

Mycobacterium tuberculosis partial rpoB gene for RNA polymerase

beta subunit, isolate 2540-97.

ACCESSION

AJ318815

VERSION

RNA polymerase beta subunit; rpoB gene.

KEYWORDS

Mycobacterium tuberculosis.

SOURCE

ORGANISM

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium

tuberculosis complex.

REFERENCE 1

Herrera, L., Jimenez, M.S. and Saez, J.A.

Molecular analysis of rifampin-resistant Mycobacterium tuberculosis

isolated in Spain (1996-2001). Description of new alleles into rpoB

gene and review

Unpublished

REFERENCE 2 (bases 1 to 618)

Herrera, L.

Direct Submission

Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro

gene	/db_xref="taxon:1773"	
	1. .618	
CDS	/gene="rpoB"	
	<1. .>618	
	/gene="rpoB"	
	/codon_start=1	
	/transl_table=11	
	/product="RNA polymerase beta subunit"	
	/protein_id="CAC87034.1"	
	/db_xref="GI:22208411"	
	/translation="TDEALLDIYRKLPGCEPTTKSAQTLLLENFFKEXRYDLARVGR	
	YKVNKKGLHVGEPITSTLTEDVVATIEYLVRHGEQTMTVPVGGVPEVETDID	
	HFGNRLRTVGLIQIRVGRSMRVVRRMTQDVEAITPQTLLIRPVVAAIKE	
	FFGTGSPSOFMGONPLSLGTHKRRLSALPGGLSRERAGLEVRD"	
BASE COUNT	124 a	192 c 207 g 95 t
ORIGIN		
Query Match	58.3%	Score 124.8; DB 1; Length 618;
Best local Similarity	77.6%	Pred. No. 4.9e-12;
Matches 166; Conservative	0; Mismatches 42; Indels 6; Gaps 1;	
QY	1	TCAAGGAGACGGCTACGACCTGGCCGCGTGGCCGGTACAGGTCAACAAGACTGG 60
Db	95	TCAAGGAGACGGCTACGACCTGGCCGCGTGGCTGCTATAAGGTCAACAAGACTGG 154
QY	61	GTCTTGGCGGTGCCAACCCCGCTCTGTGTGACTGCCACACGCTCACCGGAGAGACGTGG 120
Db	155	GGCTGCATGTCGGGAGCC-----CATCAGCTGCTGCAGCGTCGCCGAAGAAGACGTGG 208
QY	121	TGCGCACCATCGGTACTTGGTGGCTGTGCACGAGGCGCCAGACGATGACCGCCCGCG 180
Db	209	TGGCCACCATCGAATATCTGTCCGCTTGCCAGGAGTCCAGACCATGACCGTTCGCG 268
QY	181	CGGCGCTCGAGGTCGCGGTTCGAGGTTCGACGACAT 214
Db	269	CGGCGCTCGAGGTCGCGGTTCGGAACCGACGACAT 302
RESULT 7		
MTU318814	633 bp	DNA circular BCT 09-AUG-2002
LOCUS		
DEFINITION	Mycobacterium tuberculosis partial rpoB gene for RNA polymerase	
	beta subunit, isolate 1058-97.	
ACCESSION	AJ318814	
VERSION	AJ318814.1 GI:22208404	
KEYWORDS	RNA polymerase beta subunit; rpoB gene.	
SOURCE	Mycobacterium tuberculosis.	
ORGANISM	Mycobacterium tuberculosis.	
REFERENCE	1	
AUTHORS	Herrera,L., Jimenez,M.S. and Saez,J.A.	
TITLE	Molecular analysis of rifampin-resistant Mycobacterium tuberculosis	
	isolated in Spain (1996-2001). Description of new alleles into rpoB	
JOURNAL	gene and review	
REFERENCE	2 (bases 1 to 633)	
AUTHORS	Herrera,L.	
TITLE	Direct Submission	
JOURNAL	Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro	
	Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,	
	Majadahonda, Madrid. 28220, SPAIN	
FEATURES	Location/Qualifiers	
source	1. .633	
	/organism="Mycobacterium tuberculosis"	
	/isolate="1058-97"	
	/db_xref="taxon:1773"	
gene	1. .633	
	/gene="rpoB"	
CDS	<1. .>633	
	/gene="rpoB"	
	/codon_start=1	

```
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87031.1"
/db_xref="GI:22208405"
/translation="DNTVGTDEALLDIYKRLRPSPKESAOQTLELNFKEKRYDL
ARVGRYKVNKKLGLHVEPITSTLDEEDVATIEYLVRLHEGQTTMTVPGVPEVPE
TDDIDHFGNRRRLTVGELIQNIRVGMRSRMRVVRMTTQDVEAITPQTLINIRPVV
AAIKEFFGTSQLSQFMDQNNPLSLGTHRRMFALPGGLSRERAGLEVDRVPSH"
BASE COUNT      129 a   195 c   210 g   99 t
ORIGIN

Query Match      58.3%; Score 124.8; DB 1; Length 633;
Best Local Similarity 77.6%; Pred. No. 4.9e-12;
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 1 TCAAGGAGAACGCTACGACCTGGCCCGGTGGCGGTACAGGTGAACAAGAAGCTCG 60
|||||
DB 110 TCAAGGAGAACGCTACGACCTGGCCCGGTGGCGGTACAGGTGAACAAGAAGCTCG 169
|||||

QY 61 GTCTTGGCGGTGCACACCGCGTCTGGTGACTGCCACACCGTCCACCGAGGAGACGTCG 120
|||||
DB 170 GGCTGCATGTGCGCGAGCC-----CATCACGCTCGTCGACGCTGACCGAAGAAGCTCG 223
|||||

QY 121 TCGCCACCATCGGTACCTGGTGGCGCTGCACGAGGGCCAGACGATGACCGCCCCCG 180
|||||
DB 224 TGCCACCATCGAATATCTGGTCCGCTTCACGAGGGTCAGACACGATGACCGTTCGCG 283
|||||

QY 181 CGGCGCTCGAGGTCCGCGTCCGAGGTCGACGACAT 214
|||||
DB 284 CGGCGCTCGAGGTGCGCGTGGAAACGACGACAT 317
|||||

RESULT 8
MTU318816
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1255-98.
ACCESSION
AJ318816
VERSION
AJ318816.1 GI:22208408
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
Unpublished
2 (bases 1 to 637)
Herrera, L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN
FEATURES
Source
1..637
/organism="Mycobacterium tuberculosis"
/isolate="1255-98"
/db_xref="taxon:1773"
1..637
/gene="rpoB"
<1..>637
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87033.1"
/db_xref="GI:22208409"
/translacion="VTDEALLDIYKRLRPSPKESAOQTLELNFKEKRYDLARY
GRYKVNKKLGLHVEPITSTLDEEDVATIEYLVRLHEGQTTMTVPGVPEVPTDD
IDHFGNRRRLTVGELIQNIRVGMRSRMRVVRMTTQDVEAITPQTLINIRPAVAI
KEFFGTSQLSQFMDQNNPLSLGTHRRMFALPGGLSRERAGLEVDRVPSH"
BASE COUNT      126 a   202 c   212 g   99 t
ORIGIN

Query Match      58.3%; Score 124.8; DB 1; Length 639;
Best Local Similarity 77.6%; Pred. No. 4.9e-12;
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 1 TCAAGGAGAACGCTACGACCTGGCCCGGTGGCGGTACAGGTGAACAAGAAGCTCG 60
|||||
DB 101 TCAAGGAGAACGCTACGACCTGGCCCGGTGGCGGTACAGGTGAACAAGAAGCTCG 160
|||||

QY 61 GTCTTGGCGGTGCACACCGCGTCTGGTGACTGCCACACCGTCCACCGAGGAGACGTCG 120
|||||
DB 161 GGCTGCATGTGCGCGAGCC-----CATCACGCTCGTCGACGCTGACCGAAGAAGCTCG 214
|||||

QY 121 TCGCCACCATCGGTACCTGGTGGCGCTGCACGAGGGCCAGACGATGACCGCCCCCG 180
|||||
DB 215 TGCCACCATCGAATATCTGGTCCGCTTCACGAGGGTCAGACACGATGACCGTTCGCG 274
|||||

QY 181 CGGCGCTCGAGGTCCGCGTCCGAGGTCGACGACAT 214
|||||
DB 275 CGGCGCTCGAGGTGCGCGTGGAAACGACGACAT 308
|||||

RESULT 9
MTU318821
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1071-98.
ACCESSION
AJ318821
VERSION
AJ318821.1 GI:22208418
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
Unpublished
2 (bases 1 to 639)
Herrera, L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN
FEATURES
Source
1..639
/organism="Mycobacterium tuberculosis"
/isolate="1071-98"
/db_xref="taxon:1773"
1..639
/gene="rpoB"
<1..>639
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87038.1"
/db_xref="GI:22208419"
/translacion="RTDEALLDIYKRLRPSPKESAOQTLELNFKEKRYDLARY
GRYKVNKKLGLHVEPITSTLDEEDVATIEYLVRLHEGQTTMTVPGVPEVPTDD
IDHFGNRRRLTVGELIQNIRVGMRSRMRVVRMTTQDVEAITPQTLINIRPAVAI
KEFFGTSQLSQFMDQNNPLSLGTHRRMFALPGGLSRERAGLEVDRVPSH"
BASE COUNT      126 a   202 c   212 g   99 t
ORIGIN

Query Match      58.3%; Score 124.8; DB 1; Length 639;
Best Local Similarity 77.6%; Pred. No. 4.9e-12;
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 1 TCAAGGAGAACGCTACGACCTGGCCCGGTGGCGGTACAGGTGAACAAGAAGCTCG 60
|||||
DB 101 TCAAGGAGAACGCTACGACCTGGCCCGGTGGCGGTACAGGTGAACAAGAAGCTCG 160
|||||

QY 61 GTCTTGGCGGTGCACACCGCGTCTGGTGACTGCCACACCGTCCACCGAGGAGACGTCG 120
|||||
DB 161 GGCTGCATGTGCGCGAGCC-----CATCACGCTCGTCGACGCTGACCGAAGAAGCTCG 214
|||||

QY 121 TCGCCACCATCGGTACCTGGTGGCGCTGCACGAGGGCCAGACGATGACCGCCCCCG 180
|||||
DB 215 TGCCACCATCGAATATCTGGTCCGCTTCACGAGGGTCAGACACGATGACCGTTCGCG 274
|||||

QY 181 CGGCGCTCGAGGTCCGCGTCCGAGGTCGACGACAT 214
|||||
DB 275 CGGCGCTCGAGGTGCGCGTGGAAACGACGACAT 308
|||||

RESULT 9
MTU318821
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1071-98.
ACCESSION
AJ318821
VERSION
AJ318821.1 GI:22208418
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
Unpublished
2 (bases 1 to 639)
Herrera, L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN
FEATURES
Source
1..639
/organism="Mycobacterium tuberculosis"
/isolate="1071-98"
/db_xref="taxon:1773"
1..639
/gene="rpoB"
<1..>639
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87038.1"
/db_xref="GI:22208419"
/translacion="RTDEALLDIYKRLRPSPKESAOQTLELNFKEKRYDLARY
GRYKVNKKLGLHVEPITSTLDEEDVATIEYLVRLHEGQTTMTVPGVPEVPTDD
IDHFGNRRRLTVGELIQNIRVGMRSRMRVVRMTTQDVEAITPQTLINIRPAVAI
KEFFGTSQLSQFMDQNNPLSLGTHRRMFALPGGLSRERAGLEVDRVPSH"
BASE COUNT      126 a   202 c   212 g   99 t
ORIGIN

Query Match      58.3%; Score 124.8; DB 1; Length 639;
Best Local Similarity 77.6%; Pred. No. 4.9e-12;
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 1 TCAAGGAGAACGCTACGACCTGGCCCGGTGGCGGTACAGGTGAACAAGAAGCTCG 60
|||||
DB 110 TCAAGGAGAACGCTACGACCTGGCCCGGTGGCGGTACAGGTGAACAAGAAGCTCG 169
|||||

QY 61 GTCTTGGCGGTGCACACCGCGTCTGGTGACTGCCACACCGTCCACCGAGGAGACGTCG 120
|||||
DB 170 GGCTGCATGTGCGCGAGCC-----CATCACGCTCGTCGACGCTGACCGAAGAAGCTCG 223
|||||

QY 121 TCGCCACCATCGGTACCTGGTGGCGCTGCACGAGGGCCAGACGATGACCGCCCCCG 180
|||||
DB 224 TGCCACCATCGAATATCTGGTCCGCTTCACGAGGGTCAGACACGATGACCGTTCGCG 283
|||||

QY 181 CGGCGCTCGAGGTCCGCGTCCGAGGTCGACGACAT 214
|||||
DB 284 CGGCGCTCGAGGTGCGCGTGGAAACGACGACAT 317
|||||

RESULT 8
MTU318816
LOCUS
DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1255-98.
ACCESSION
AJ318816
VERSION
AJ318816.1 GI:22208408
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE
1
Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
Unpublished
2 (bases 1 to 637)
Herrera, L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN
FEATURES
Source
1..637
/organism="Mycobacterium tuberculosis"
/isolate="1255-98"
/db_xref="taxon:1773"
1..637
/gene="rpoB"
<1..>637
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87033.1"
/db_xref="GI:22208409"
/translacion="VTDEALLDIYKRLRPSPKESAOQTLELNFKEKRYDLARY
GRYKVNKKLGLHVEPITSTLDEEDVATIEYLVRLHEGQTTMTVPGVPEVPTDD
IDHFGNRRRLTVGELIQNIRVGMRSRMRVVRMTTQDVEAITPQTLINIRPAVAI
KEFFGTSQLSQFMDQNNPLSLGTHRRMFALPGGLSRERAGLEVDRVPSH"
BASE COUNT      129 a   195 c   210 g   99 t
ORIGIN
```

Best Local Similarity 77.6%; Pred. No. 4.9e-12;									
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;									
QY	1	TC AAGGAGAAGCGCTACGACCTGGCCCGCTGGGCGGTACAAGGTGAACAAGAAGCTGG	60						
Db	101	TC AAGGAGAAGCGCTACGACCTGGCCCGCTGGGCGGTACAAGGTGAACAAGAAGCTGG	160						
QY	61	GTCTTTGGCGGTGCCAACCCGGCTCTGGTGACTGCCACACCGCTACCGAGGAAGCTCG	120						
Db	161	GGCTGCATCTCGCGAGCC-----CATCAGCTGCTGACGCTGACCGAAGAAGACGCTG	214						
QY	121	TGCCCACCATCGGTTACCTGGTGGCGCTGACGAGGGCCAGACACGATGACCGCCCGG	180						
Db	215	TGCCCACCATCGAATATCTGGTCGCTTGGACGAGGTCAGACACGATGACCGTTCCGG	274						
QY	181	CGGGCTCGAGGTCGGGTGCGAGGTCGACGACAT	214						
Db	275	CGGGCTCGAGGTCGGGTGGAACCCAGCAGAT	308						
RESULT 10									
I50706									
LOCUS	I50706		970 bp	DNA	linear	PAT 07-OCT-1997			
DEFINITION	Sequence 1 from patent US 5643723.								
ACCESSION	I50706								
VERSION	I50706.1 GI:2472409								
KEYWORDS	.								
SOURCE	Unknown.								
ORGANISM	Unclassified.								
REFERENCE	1 (bases 1 to 970)								
AUTHORS	Pensing,D.H., Hunt,J.J., Young,K.K.Y., Felmlee,T.A., Roberts,G.D. and Whelan,A.Christian.								
TITLE	Detection of a genetic locus encoding resistance to rifampin in mycobacterial cultures and in clinical specimens								
JOURNAL	Patent: US 5643723-A 1 01-JUL-1997;								
FEATURES	Location/Qualifiers								
source	1..970								
BASE COUNT	182 a 302 c 330 g 156 t								
ORIGIN									
Query Match 58.3%; Score 124.8; DB 6; Length 970;									
Best Local Similarity 77.6%; Pred. No. 4.6e-12;									
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;									
QY	1	TC AAGGAGAAGCGCTACGACCTGGCCCGCTGGGCGGTACAAGGTGAACAAGAAGCTGG	60						
Db	26	TC AAGGAGAAGCGCTACGACCTGGCCCGCTGGGCGGTACAAGGTGAACAAGAAGCTGG	85						
QY	61	GTCTTTGGCGGTGCCAACCCGGCTCTGGTGACTGCCACACCGCTACCGAGGAAGCTCG	120						
Db	86	GGCTGCATCTCGCGAGCC-----CATCAGCTGCTGACGCTGACCGAAGAAGCTCG	139						
QY	121	TGCCCACCATCGGTTACCTGGTGGCGCTGACGAGGGCCAGACACGATGACCGCCCGG	180						
Db	140	TGCCCACCATCGAATATCTGGTCGCTTGGACGAGGTCAGACACGATGACCGTTCCGG	199						
QY	181	CGGGCTCGAGGTCGGGTGCGAGGTCGACGACAT	214						
Db	200	CGGGCTCGAGGTCGGGTGGAACCCAGCAGAT	233						
RESULT 11									
AX111339									
LOCUS	AX111339		3534 bp	DNA	linear	PAT 30-APR-2001			
DEFINITION	Sequence 2072 from Patent W00123604.								
ACCESSION	AX111339								
VERSION	AX111339.1 GI:13927631								
KEYWORDS	.								
SOURCE	Mycobacterium tuberculosis.								
ORGANISM	Mycobacterium tuberculosis Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;								

```
/product="RNA-polymerase beta subunit"
/protein_id="AAA20242.2"
/db_xref="GI:7144499"
/translation="MLEGILADRSQSKTAASPSRPOSSNNNSVPCAPNRVVSFAKL
REPLEVGLDQDTDFEWLIGSPRWSAERGDVNPVGGLEEVLYELSELPIEDFSGS
MSLFSDFRFDVKAQVDECKDKMTYAAPLFTVAEFNNNTGEIKSOTVPMGDFPMH
TEKGTPIINGTERRVVSQIVRFGVFDETIDKSTDKTLSHVKIPSGRWALFEDVDK
RDTGVVIDDKRRQPVVLLKALGWTSEQIVERFGFSEIMRSTLEKONTGWTDEALID
IYKLRPGEPPTKESAOQLLENLFFKERYDLARVGRVKNKGLGLHVGEPITSTSLT
EEDVATIEYLRHLEGQTITVPGVEVPVETDDIDHFGNRRRTVCELONQIRVG
MSNRVERHMTQDVEAITQTQLINIRPVYAAIKKEFGTSQLSOFMDONPLSLGLT
HKRRLSALPGGLSRRERAGLEVDRVHPSHYGMCPETIEPGNIGLIGLSLYARVNP
FGRIETPKYVDVGVSDIVLTADAEDRHVYAQANSPIDADGRFVPEPRVLRKAG
LVGVPSSEVDVSPROMVSATAMIPFLEHDDANRALMGANNORQAVPLVRSEAP
EYTGEMELRAAIDAATSSQESGVIEVSADYITVMHNGTTRTYMRKFAFARNHGT
EYTGEMELRAAIDAATSSQESGVIEVSADYITVMHNGTTRTYMRKFAFARNHGT
ANOCPLVDAGDRVEAGOVADGCTDDGEMALGNLLVAIMPWEHNGEDAIILSNRL
VEEDVLTSIHIEHEIDARDTKLGAEEITRDIPNISDEVLDLDERGIVRIGAEVRDG
DILVGVTPKGETELPBERLLRAIFGEKAREVRTSLKVPHGSEKVGIRVFSRED
EDELPAVGNELVRVYVAQRKISDGDKLAGRNGKGVIGKILPVEDMPLADGTPVDI
ILNTHGVPRRMNIGQILETHLGCASHGKWDAAKGVDPDWAARLPDELLERQPNATIS
TPVFDGAQAEALQGLSCTLPNRDGLVDADGKAMLFDRSGSEFPYPTVTVGYMIIM
KLHLVDDKIHARSTGYSNITQOPLGKAQFGGOFGEKMECWAQYGAAYTLOELL
TIKS"
```

BASE COUNT 723 a 1173 c 1293 g 664 t

Query Match 58.3%; Score 124.8; DB 1; Length 3853;
Best Local Similarity 77.6%; Pred. No. 3.9e-12;
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 1 TCAAGGAGAACGCTACACCTGGCCCGCGTGGCGGTCAAGGTCAACAAAGAAGCTGG 60
|||||
Db 1477 TCAAGGAGAACGCTACACCTGGCCCGCGTGGCGGTATAGGTCACAAAGAAGCTCG 1536
|||||

QY 61 GTCCTGGCGGTGCCAACCGCGCTCTGTGACTGCACACCGCTCACCGAGGAAGACGTGC 120
|||
Db 1537 GGTGTCATGTCGCGGAGCC-----CATCAGCTGCTGCACGCTCACCGAAGACGTCG 1590
|||

QY 121 TGCCACCATCGGTACTGTGTCGCTTGCACGAGGGCCAGCACCATGACGCCGCCCG 180
|||||
Db 1591 TGCCACCATCAATATCTGTCGCTTGCACGAGGGTCAGACCATGACCGCTTCGCG 1650
|||||

QY 181 GCGGCTCGAGTCCGCGTCGAGGTCGACGACAT 214
|||||
Db 1651 GCGGCTCGAGTCCGCGTGGAAACGACGACAT 1684
|||||

RESULT 13

MSGRPOB 5084 bp DNA linear BCT 13-SEP-1994
LOCUS Mycobacterium tuberculosis RNA polymerase beta-subunit (rpoB)
DEFINITION gene, complete cds and RNA polymerase beta'-subunit rpoC gene,
partial cds.
L27989
VERSION 1 GI:468333
KEYWORDS RNA polymerase beta-subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis (strain Rv).
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
REFERENCE 1 (bases 1 to 5084)
AUTHORS Miller,L.P., Crawford,J.T. and Shinnick,T.M.
TITLE The rpoB gene of Mycobacterium tuberculosis
JOURNAL Antimicrob. Agents Chemother. 38 (4), 805-811 (1994)
MEDLINE 94304130
PUBMED 8031050

FEATURES
source Location/Qualifiers
1..5084
/organism="Mycobacterium tuberculosis"
/strain="Rv"
/db_xref="taxon:1773"
1065..4598

gene

```
/gene="rpoB"
1065..4598
/gene="rpoB"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="RNA polymerase beta-subunit"
/protein_id="AAA21416.1"
/db_xref="GI:468334"
/translation="MLEGILADRSQSKTAASPSRPOSSNNNSVPCAPNRVVSFAKL
REPLEVGLDQDTDFEWLIGSPRWSAERGDVNPVGGLEEVLYELSELPIEDFSGS
MSLFSDFRFDVKAQVDECKDKMTYAAPLFTVAEFNNNTGEIKSOTVPMGDFPMH
TEKGTPIINGTERRVVSQIVRFGVFDETIDKSTDKTLSHVKIPSGRWALFEDVDK
RDTGVVIDDKRRQPVVLLKALGWTSEQIVERFGFSEIMRSTLEKONTGWTDEALID
IYKLRPGEPPTKESAOQLLENLFFKERYDLARVGRVKNKGLGLHVGEPITSTSLT
EEDVATIEYLRHLEGQTITVPGVEVPVETDDIDHFGNRRRTVCELONQIRVG
MSNRVERHMTQDVEAITQTQLINIRPVYAAIKKEFGTSQLSOFMDONPLSLGLT
HKRRLSALPGGLSRRERAGLEVDRVHPSHYGMCPETIEPGNIGLIGLSLYARVNP
FGRIETPKYVDVGVSDIVLTADAEDRHVYAQANSPIDADGRFVPEPRVLRKAG
LVGVPSSEVDVSPROMVSATAMIPFLEHDDANRALMGANNORQAVPLVRSEAP
EYTGEMELRAAIDAATSSQESGVIEVSADYITVMHNGTTRTYMRKFAFARNHGT
EYTGEMELRAAIDAATSSQESGVIEVSADYITVMHNGTTRTYMRKFAFARNHGT
ANOCPLVDAGDRVEAGOVADGCTDDGEMALGNLLVAIMPWEHNGEDAIILSNRL
VEEDVLTSIHIEHEIDARDTKLGAEEITRDIPNISDEVLDLDERGIVRIGAEVRDG
DILVGVTPKGETELPBERLLRAIFGEKAREVRTSLKVPHGSEKVGIRVFSRED
EDELPAVGNELVRVYVAQRKISDGDKLAGRNGKGVIGKILPVEDMPLADGTPVDI
ILNTHGVPRRMNIGQILETHLGCASHGKWDAAKGVDPDWAARLPDELLERQPNATIS
TPVFDGAQAEALQGLSCTLPNRDGLVDADGKAMLFDRSGSEFPYPTVTVGYMIIM
KLHLVDDKIHARSTGYSNITQOPLGKAQFGGOFGEKMECWAQYGAAYTLOELL
TIKSDDTVGRVKYEAIVKGENIPGPIPESEFKVLLKELQSLCLNVEVLSDDGAAIL
REGDEDELERAAANUGINLSRNESAFEDLA"
```

BASE COUNT 969 a 1534 c 1691 g 890 t

Query Match 58.3%; Score 124.8; DB 1; Length 5084;
Best Local Similarity 77.6%; Pred. No. 3.7e-12;
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;

QY 1 TCAAGGAGAACGCTACACCTGGCCCGCGTGGCGGTCAAGGTCAACAAAGAAGCTGG 60
|||||
Db 1966 TCAAGGAGAACGCTACACCTGGCCCGCGTGGCGGTATAGGTCACAAAGAAGCTCG 2025
|||||

QY 61 GTCCTGGCGGTGCCAACCGCGCTCTGTGACTGCCACCGCTCACCGAGGAAGACGTCG 120
|||
Db 2026 GGTGTCATGTCGCGGAGCC-----CATCAGCTGCTGCAGCTGACCGAAGACGTCG 2079
|||

QY 121 TGCCACCATCGGTACTGTGTCGCTTGCACGAGGGCCAGCACCATGACGCCGCCCG 180
|||||
Db 2080 TGCCACCATCAATATCTGTCGCTTGCACGAGGGTCAGACCATGACCGCTTCGCG 2139
|||||

QY 181 GCGGCTCGAGTCCGCGTCGAGGTCGACGACAT 214
|||||
Db 2140 GCGGCTCGAGGTCGCGGTGGAAACCGACGACAT 2173
|||||

RESULT 14

AE006964 19352 bp DNA linear BCT 27-APR-2001
LOCUS Mycobacterium tuberculosis CDC1551, section 50 of 280 of the
DEFINITION complete genome.
ACCESSION AE006964 AE000516
VERSION AE006964.1 GI:13880217

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE

JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
FEATURES
source

gene
CDS

gene
CDS

gene
CDS

gene
CDS

gene
CDS

/translation="MLDVNFDELRLGLATAEDIRQWSYGEVKKPETINYBTLKPEXD
GLFCRIFGPTRWECYCGKTKRFRFKGLICERCVCVEVTRAKVRERMRHGLHLEAPVT
HIWYFKGPSRLGYLLDLAPKDLKLEKIYFAAYVITSVDEMRHMLSTLEAMAVERK
AVEDQDGEARAKLELAELAEAGAKADARKVRDGGERRMQRIQDRAQRELD
LEDIWSFTKLPKOLIVDENLYRELVDYGEYFTGAMGAESIOLKLIENFIDIAEAS
LRDVRNGKQKKRLARLKVAAFOQSGNSPMGMVLDAVPVIPPPELRPMVQDGG
FATSDNDLYRRVINNRHLKRLIDLGAPELIIVNKKPKLQESVDLFDNGRRGPVT
GPGNRLKSLDLKQGRFRQNLGKRVDYSGRSVIVGFGPLKLOCCGFLPKMALE
LFKPFVMKRLVDLNAHQNKSKRMVRFORFQVMDVLEEVIAEHVLLNRAPTLHRIG
IQAFPEMLVEGKAIQHLPLVCEAFNAFEDQGMVHLPLSAEAQAARILMLSSNII
SPASGRPLAMPRLDMVTGLYLTTEVPEDTGEYQPSAGDHPETGYVSSPAAINAAIR
GVSVRAKIKVLTOLRPPVEIEAEELFGHSGMGDAMMAETTLGVRVNFNELLPLGYR
FVKNOMHKKVQAAIINDLAERYPMIVVAQTVDKLDAGFYHATSEVTVSNADVLPP
RKKEILDHYERADKVEKQFQRCALNHDNERNEALVEIKKEATDEVGQALREHPDDNP
IITIVDSGATGNFTQTRTLAGMKGLVTPKGEFIPRPKSSGFLEVTLEYVFINTHGA
RKLADATLRTADSGYLTRLVVDVSDQVIVREHDCQTERGIVVELAERAPDGTILRDP
YLETSAYARTLTGTDVADAGNVIVVERQDLGDPEIDALLAAGITQKVRSVLTCTATST
GVCACTGRSMATGKLVLDIGEAAGVIAAQSIGEPGTQTMRTFHGGVGEDITGGLPR
VOELFEARVPKGAPIADVTGRVLEDGEREYKJITIVDDGSEEVYDKISKRORLRV
FKHEDGSESVLSDGHRVEVGOOLMEGSDAPHEVLAVQSPREVQHLHREVOEYRAQG
VSIHDKHIEVIVRQMLRVTIIDSGSTFELPGLSLIDRAEFKAENRRVVAEGEPPAAGR
PVLGGITKASLATDSMLSAASQFETTRVLTDAAINCRSKLNGKENYIIGKLIIPAGT
GINRYRNTAVOPTEEARAAAYTIPSYEDQYTSDFGAATGAAPVLDYDYSYDR"
/gene="WT0697"
complement(7691..8065)
/gene="WT0697"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK44923.1"
/db_xref="GI:13880220"
/translation="MFDAAAITNPGHAWASAMERSGLLECVAGLEQPFGEFTADKL
NPDGSSRRVPRRQADGGIATHVERGGQSGQGVVPMHGFALAMQDRLLIHH
GQTONRFAQAPRVRCVCSPT"
complement(8058..9972)
/gene="WT0698"
/note="this region contains an authentic frame shift and
is not the result of a sequencing artifact; identified by
Glimmer2; putative; conserved hypothetical protein,
authentic frameshift"
10167..10925
/gene="WT0699"
10167..10925
/note="identified by match to PFAM protein family HMM
PF011261"
/codon_start=1
/transl_table=11
/product="AP endonuclease, family 2"
/protein_id="AAK44924.1"
/db_xref="GI:13880221"
/translation="MLIGSHVSPDPLAAAEAGADVQVIFLGNPQSWKAPKPRDDAA
AKAATLPIYHAPYLINLASANNVRIPSRKILQETCAAAADICAAAVIVHGGHVA
DNDIDKQFORWKALDRLETEVPVYLENTAGDGHAKAREFTIARLMDVIGDTGIGFC
LTCVHTWAGEALTDVADRIKAITCIDRLVHCNDSRDEAGSGDRHANLGSQIDPDL
LVAAKAGAPVICTADQGRKDDIAFLRERTGS"
10957..11799
/gene="WT0700"
10957..11799
/gene="WT0700"
/note="similar to GB:U00012 PID:466863; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hydrolase/esterase, putative"
/protein_id="AAK44925.1"
/db_xref="GI:13880222"
/translation="MLRRVAILLAVALFAGSCGTRLAAGFGNSVHTLDVDCAGR
SVRLYKPGCLSPSAPLVMLHGGFGSAKQERSYGMDELASEKELVAYIVHGHVAD
ANGGGCCRPAREGVDDIGFVRVYADIANVYSIDPARVYVTGMSNGAIVMSYTLACNT
SIFAAIGVVSQTLDPQCSRPVSVIHITGTADPLVRYHGGPGAGFARIDGPPVVDLN

Mycobacterium tuberculosis CDC1551.
Mycobacterium tuberculosis CDC1551
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium tuberculosis complex.
1 (bases 1 to 19352)
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O.,
Peterson, J., Deboy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E.,
Kolony, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M.,
Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H.,
Gill, J., Mikula, A. and Bishai, W.
Whole genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains
Unpublished
2 (bases 1 to 19352)
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O.,
Peterson, J., Deboy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E.,
Kolony, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M.,
Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H.,
Gill, J., Mikula, A. and Bishai, W.
Direct Submission
Submitted (25-APR-2001) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
1..19352
/organism="Mycobacterium tuberculosis CDC1551"
/strain="CDC1551"
/db_xref="taxon:8331"
/note="clinical strain"
163..3699
/gene="WT0695"
163..3699
/gene="WT0695"
/note="similar to GB:L27989 GB:L05910 GB:U12205 SP:P47766
PID:149992; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA-directed RNA polymerase, beta subunit"
/protein_id="AAK44921.1"
/db_xref="GI:13880218"
/translation="MLEGCIILADSRQSKTAASPSRQSSNSNVPGAPNRVYSFAKL
REPLEVGLLDVQTFSEWILIGSPRMRSAERGDPNPGGLEVEIYELSPIDFSGS
MSLSFGDPDVKAPVDECKDMYAAPLFVTAEFINNTEGIEKSTQMDQNNPLSGIT
TEKGTFTINTERVVSQVLRSPCVFDEITDKTDKTLHSVKVIPSRAWLEFDYDK
RDTGVDIRKRRQPVYLLKALNTSEQIVREPGSEIMRSLKDNVGTDEALLD
LYRLRGEPTRESAQTLLENLFFREKRYDLARVGRVKNKKLGLHVGEPITSSLT
EEDVATYELVRLHEGOTTMTVPGGVEVPVETDDIDHFNRLRTVGEILQNIQIRVG
MSRMERVVERMTQDVEAITPQTLINIRPVAAIKFEFTQSLSQFMDQNNPLSGIT
HKRLSALPGLSRRERAGLEVDRVHPSHYRCMPTETPEGNIGLISLVSRYRNP
EGFLETYRKYVDGVSDIEVYLTADEEDRHHVAAQANSPIADGRFVEPRLVRKAG
EVEYVPSVEDYHDVSPQWVSATANIPELHDDANRALGMANQOAVPLVRSEAP
LVGTGMLERAIDAGVVAAEESGVEIEVSADYITVHNDGTRRTYMRKFAKSNHGT
CANQCPIDVADGRVAGQVADGCTDDGEMALGNLLVAIMPWEHNYEDAILNSR
LVEDVLTSTHIEHEIDARTDLGAEETRIDIPNIDEVLADLDRGIVRIGAEYRD
GDILVGKTPKGETPPEERLLRAIPGKAREVROTSLKVPHGSGKIGIVFVSRE
DEDELPAVNELRVVYVQAKRISDQGLAGRGNKGVIKILPVEDMPELADGTVD
IILNTHGVPRMNIQOILETHLGNCAHSCHKVDAKGVDPDRAARLDELLEAOPNAY
STPVFGAELAGLGLSCLTPNRRDGLVDADKAMLPDRSGSEFPFPYVGYMYI
MKLHLVDDKIHARSOTGPYSMITQOPLGGKAQFGQRFGEKMECWAQYGAAYTLQEL
LTKSDDTGVRKYVEAIVKGENIPEGIPESFKVLLKELQSLCLNVEVLSDDGAIE
LRECEDELERAAANILNLSRNESASVEDLA"
3744..7694
/gene="WT0696"
3744..7694
/gene="WT0696"
/note="similar
similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA-directed RNA polymerase, beta-prime subunit"
/protein_id="AAK44922.1"
/db_xref="GI:13880219"


```
AFWREYNRCGALDTTTTEGVPVTTSGATCADNRVRVLLTVDDAGHRWPSFATQTLWRFFA
AHER"
11859..11487
/gene="MT0701"
11859..11487
/gene="MT0701"
/notes="similar to SP:P33224 GB:L20915 PID:457172
PID:457174 PID:537028; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="acyl-CoA dehydrogenase, putative"
/protein_id="AAK44926.1"
/db_xref="GI:13880223"
/translation="MSDTHVTVNQVPLENYNPASPVLIEALIOEGQWGLDEVNEY
GAISACQORGEALNRNPILH7HDAYRYRVEVDEYDPAYHELMRTAITHGMHAAP
WADRRPGARVGAELAKTSVTVPEGHICPISMTYAVVPALRYNSELAAYVEPLITSREY
DELPAPATTKAGTITAGMSMTEKGGSDVRAGTTOATPNADGSYSLTGHKWTSAVPMCD
IFLVLAQAPDGLSCFLPRVLPDGTNRNMFLORLKDLGNHANASSEVEYDGAVALV
GEGRGVPTLIEVNLTRLDLCALGSATSMRTGLTRAVHHAQHRKAFAYLIDQPLMRN
VLADLAVERAATIVAMRNAGATDNARVGNTEALLRIGLAAAYWCKRSTAHAAE
ALECUGNGYVSDSGMPRLRYREAPLMGWEGSGNVSALDTLRMATRPACVEVLFDL
ARSAQDPLRDLRVERLRPQLGLDITGYRARKIAGDIEDICLALQGSLLVRHGHPAVAEA
FLATRLGGQWGGAYGTMPPAGLDLAPILERALVKG"
13498..14436
/gene="MT0702"
13498..14436
/gene="MT0702"
/notes="similar to GP:3885480; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="enoyl-CoA hydratase/isomerase family protein"
/protein_id="AAK44927.1"
/db_xref="GI:13880224"
/translation="MTHAIRPVDLNTMTYEVTRIGRIATFNRPKGNALIIDPTPL
ELSAVERADLPDGVHVIIVSGRGFCAGFDLSAAGSSSTGGGAYQGVLDGKT
QAVNLHPNPWPMIDYQNMRSFVRGFASLMHADRPYVKIHGYCVAGGTDLALHADQ
VTAAADAKGYPTPRVWGPVPAAGLWHRLDQORAKRLFTDTCITGAQAEMGLAVEA
PEPADUDTERLAKIAALPVNQLIMVKALNSALLOOGVATSRMSTVDFDGAARHT
PGHAFVADAVEHGFRDAVRRDEPFGDYGROASRV"
14439..15161
/gene="MT0703"
14439..15161
/gene="MT0703"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
Query Match 58.3%; Score 124.8; DB 1; Length 19352;
Best Local Similarity 77.8%; Pred. No. 3.1e-12;
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;
Qy 1 TCAAGGAGAGCGCTACGACCTGGCCCGCTGGCGCGCTACAAGGTGAACAAGACTCG 60
Db 1064 TCAAGGAGAGCGCTAGACCTGGCCCGCTCGTGCCTATAGGTCACAGAAGCTCG 1123
Qy 61 GTCTTTGGCGGTGCCAACCCGGCTCTGGTGACTGCCACACCGCTCACCGAAGAGCTCG 120
Db 1124 GGCTGTCATTCGGCGAGCC-----CATCACGTGCTGCGACGCTGACCGAAGAGAGCTCG 1177
Qy 121 TCGCCACCATCGGCTACCTGGTGGCTGCACCGAGGCCACACGATGACCGCCCCCG 180
Db 1178 TGGCCACCATCGGAATATCTGGTCCGTTGTCACGAGGGTCAGACCAGATGACCGTTCCCG 1237
Qy 181 GCGGCCCTCGAGGTCCCGGTCAGGTCGACGACAT 214
Db 1238 GCGCGCTCGAGGTGCCGGTGAACACGACGACAT 1271
RESULT 15
MTCI376
```

```
LOCUS MTCI376 19770 bp DNA linear BCT 03-AUG-2001
DEFINITION Mycobacterium tuberculosis H37Rv complete genome; segment 32/162.
ACCESSION Z95972 ALI23456
VERSION Z95972.1 GI:3261790
KEYWORDS
SOURCE Mycobacterium tuberculosis H37Rv.
ORGANISM Mycobacterium tuberculosis H37Rv
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
REFERENCE 1 (bases 1 to 19770)
AUTHORS Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,
Harris,D., Gordon,S.V., Eiglmeier,K., Gas,S., Barry III,C.E.,
Tekoa,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,
Connor,R., Davies,R., Devlin,K., Feltwell,T., Gentles,S.,
Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J.,
Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A.,
Rajandream,M.A., Rogers,J., Rutter,S., Seeger,K., Skelton,S.,
Squares,S., Squires,R., Sulston,J.E., Taylor,K., Whitehead,S. and
Barrell,B.G.
DECIPHERING the biology of Mycobacterium tuberculosis from the
complete genome sequence
NATURE 393 (6685), 537-544 (1998)
98295987
PUBMED 9634230
2 (bases 1 to 19770)
DIRECT Submission
Parkhill,J.
Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 27, 1998 this sequence version replaced gi:2143285.
NOTES:
Details of M. tuberculosis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have
been renumbered from the original cosmid submissions but the old
gene designations are in brackets after the new gene numbers.
Gene prediction was based on a Hidden Markov Model of TB genes
implemented in TBparse (Krogh) supplemented with visual inspection
of positional base preference in codons, especially where there is
an increase in the observed/expected third position G + C.
CAUTION: In some cases we may not have predicted the correct
initiation codon. Where possible we choose an initiation codon
(atg, gtg, or ttg) which is preceded by an upstream ribosome
binding site sequence (optimally 5-13bp before the initiation
codon). If this cannot be identified we choose the most upstream
initiation codon.
Location/Qualifiers
1..19770
/organism="Mycobacterium tuberculosis H37Rv"
/strain="H37Rv"
/db_xref="taxon:83332"
/clone="I376"
57..61
/notes="agga, possible rbs upstream of Rv0654"
68..1573
/gene="Rv0654"
68..1573
/gene="Rv0654"
/notes="Rv0654", (MTCI376.22), len: 501. unknown, FASTA
score: Q53353 LIGNOSTILBENE-ALPHA, BETA-DIOXYGENASE (485
aa) opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity
in 523 aaoverlap). Also similar to M. tuberculosis protein
MTCY21C12.07C (29.5% identity in 522 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0654"
/protein_id="CAB09380.1"
/db_xref="GI:2143307"
/db_xref="SPTREMBL:O06785"
```

/translation="MTTAAQAAEQNPYLEGLAPVSTEVATDLPVTGRIPEHLGRY
LRNPNVAVDPATVHWFTGDAMVHGVALRGDKARWYRNWRVTPAVCAALGEPISA
RPHPTGIIEGSGNTVNLVHAGRTLALVAGVYNYELTDELDTVGCDFDTLHGYYT
AHQDPHTGELHNAVSTASARGHRVQISVIGTDGHARTVDIEVAGSPMHHSFSLTDN
YVYIDLPTVTFPMQVVPASVPRMQRPARLVQSLGRVRIIPDPDIAALGNRMOGHSN
RLPYAMPNPSPARVPGREGNEDVRFWDIEPCVYVHPLNAYSECRNGAEVLVDVV
YRSMEDRRRGDSRPSLDRWTINLATGAVTAECDRDRAGREFRINETLVGGPHR
FAVTVTEGFLVAGAAALSTPLYKDCVTSSTVASLDPLDILIGEMVFPVNP SARAE
DDGILMCGWHRGRDEGQLLLLDQAQTLESIAIVHLPRQVPMGFHGNWAPT"
1585. 2664
/gene="Rv0655"
1585. 2664
/gene="Rv0655"
FASTA score: YRBF_ECOLI P45393 hypothetical abc
transporter atp-binding (269 aa), opt: 644 z-score: 721.8
E(): 3.4e-33 (38.5% identity in 244 aa overlap); contains
PS00017 ATP/GTP-binding site motif A, PS00211 ABC
transporters family signature, highly similar to M. leprae
MKL_MYCLE P30769 possible ribonucleotide transport atp-
(347 aa) opt: 2021 z-score:2244.4 E(): 0. (92.2% identity
in 335 aa overlap). Also similar to many other M.
tuberculosis ABC transporters eg. MTCY253.24 (33.6%
identity in 241 aa overlap)
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0655"
/protein_id="CAB09379.1"
/db_xref="GI:2143306"
/db_xref="SWISS-PROT:O06784"
/translation="MRYSDSVHTGWRQPRASTGFPMGVSIENVGLTKFGSRIWE
DVTLPAGEVSLVIGSGTSKVFSLKSLGLLRPERGSIIDGTIDIEGSAKELVEI
RTFLVGFODGALFGSMNLVNTAFPLRHTKKKESEIRDIVMEKALVGLGDEKFF
PEIEGGMRKAGRALVLDPQIILCEPDSDGLDPVTVLSOLINDINAQIDARTIL
VTHNINARTVPDNNMGLFRKHLVFMFGPREVLVTSDEPVVROFLNGRIGPOMSEE
KDEATMAEALDLAGHAGVVEIEGVPPOISATPCMPERKAVARQARVREMLTFL
PKQAQAAILDDEGTHKVAHVEIGQ"
1762. 1785
/gene="Rv0655"
/note="PS00017 ATP/GTP-binding site motif A"
2074. 2118
/gene="Rv0655"
/note="PS00211 ABC transporters family signature"
complement(3052. 3435)
/gene="Rv0656c"
complement(3052. 3435)
/gene="Rv0656c"
/note="Rv0656c, (MTCI376.20, unknown), len: 127 aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0656c"
/protein_id="CAB09378.1"
/db_xref="GI:2143305"
/db_xref="SPTREMBL:O06783"
/translation="MAAATTGTGHRGLELRAAORAVGSCPEPORAECRSARNADPEFQ
MSRMFGDVPDPVPKPSWVRMIDSAQHLRAGAVGALSVDLLICDTAARGLVVLH
DDADYELAEHLDPDIRVRVVSADD"
complement(3530. 3685)
/gene="Rv0657c"
complement(3530. 3685)
/gene="Rv0657c"
/note="Rv0657c, (MTCI376.19), unknown, len: 51 aa; similar
to several other M. tuberculosis hypothetical proteins eg.
YW08_MYCTU Q10848 hypothetical 8.9 kd protein cy39.08c (80
aa), fasta scores; opt: 107 z-score: 182.3 E(): 0.0036,
45.8% identity in 48 aa overlap. Also similar to MTCY48.5
andAL020958|SC4H8_7 streptomycetes coelicolor cosmid 4H8 (66
aa), 41.0% identity in 39 aa overlap."
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0657c"
/protein_id="CAB09377.1"
/db_xref="GI:2143304"

/db_xref="SPTREMBL:O06782"
/translation="MSVTOIDLDDELADVMRIAIVHKKAVNLAMRDYVERFRRIE
ALARSRE"
complement(3691. 3696)
/note="possible RBS upstream of Rv0657c"
complement(3761. 4477)
/gene="Rv0658c"
complement(3761. 4477)
/gene="Rv0658c"
/note="Rv0658c, (MTCI376.18), len: 238, unknown,
probablemembrane protein, similar to YPRB_ECOLI P33774
hypothetical24.3 kd protein (urf 1) (217 aa), fasta
scores; opt: 174 z-score: 215.8 E(): 5.3e-05, (25.6%
identity in 223 aa overlap). Also similar to MTCY359.10
(28.7% identity in 178 aa overlap)
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0658c"
/protein_id="CAB09376.1"
/db_xref="GI:2143303"
/db_xref="SPTREMBL:O06781"
/translation="MEAGRDTVAPSHRWGLGAFVLVFLVASTSLAVVLVTHGHPV
SAGVLALALAAPTVYAGLAILITRLRNGRLTDLRWSMRGLRLGLMFGFGMLVY
IPASLYTAIVGPEANSVAVRIFGVRASPMALVFLVYVFAVPLCEETIYRGLLMG
IVDRRWGRMAALVYTVVFAHALHLEFARAPLLVYVPIALAREFVSGGLASIVTHQV
TNLLPGIVLLGLTGAIPLP"
complement(4480. 4483)
/note="possible RBS upstream of Rv0658c"
complement(4753. 5061)
/gene="Rv0659c"
complement(4753. 5061)
/gene="Rv0659c"
/note="Rv0659c, (MTCI376.17), len: 102; unknown, similar
to YW28_MYCTU Q10867 hypothetical 12.3 kd protein cy39.28
(114 aa), fasta scores; opt: 144 z-score: 213.2 E():
7.3e-05, 30.8% identity in 107 aa overlap. Also similar to
MTCY09F9.22 (32.7% identity in 101 aa overlap)
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0659c"
/protein_id="CAB09387.1"
/db_xref="GI:2143302"
/db_xref="SPTREMBL:O06780"
/translation="MRGELWFAATPGGDRPVLVLRDPVADRIGAVVVALTRTRRG
LYSELELTAVENRVPSDCVNFNDNIHTLPRTAFRRITRLSPARLHEACOTLRASTGC
"
complement(5048. 5293)
/gene="Rv0660c"
complement(5048. 5293)
/gene="Rv0660c"
/note="Rv0660c, (MTCI376.16), len: 81, some similarity to
IAF016485_130 Halobacterium sp: NRC-1 plasm (100 aa),
32.4%identity in 74 aa overlap"
/codon_start=1
Query Match 58.3%; Score 124.8; DB 1; Length 19770;
Best Local Similarity 77.6%; Pred. No. 3.1e-12;
Matches 166; Conservative 0; Mismatches 42; Indels 6; Gaps 1;
QY 1 TCAAGGAGAGCGGTACGACCTGGCCGCGGTGGCGGTGACAGGTGAACAAGAAGCTGG 60
|||||
Db 10758 TCAAGGAGAGCGGTACGACCTGGCCGCGGTGGCGGTATAGGTCACAGAAGCTCG 10817
|||||
QY 61 GTCCTGGCGGTGCCAACCCCGCTCTGCTGACTGCCACACGCTCACCGAGGAAGAGCTCG 120
|||||
Db 10818 GCCTGCATGTCGGGAGGCC-----CATCACGCTGCTGCAGCGCTGACCGAAGAAGACGCTCG 10871
|||||
QY 121 TCGCCACCATCGGTACTGTGCGCTGCACGAGGCGCCAGACGATGATCACGCCGCCCG 180
|||||
Db 10872 TGGCCACCATGGAATATCTGCTCGCTTGCAGGAGGTGACACCGATGACCGTTCCGG 10931
|||||
QY 181 CGGCGCTCGAGGTCCCGGTGAGGTGCGACGACAT 214
|||||

Db 10932 GCGCGTCGAGGTCCCGGTGGAACCGACGACAT 10965

Search completed: November 13, 2002, 01:25:30
Job time : 681.07 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 12, 2002, 16:59:42 ; Search time 1109.13 Seconds
(without alignments)
3037.202 Million cell updates/sec

Title: US-09-697-123B-1
Perfect score: 208
Sequence: 1 tcaaggagagcgtacgac.....ccggttgagaccgacacat 208

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0
Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estinv:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
C 1	163.2	78.5	1282 9	AI770311 42 Mycoba
2	47	22.6	517 13	BJ233830 BJ233830
3	47	22.6	649 12	BC908023 TAlr1164F
4	47	22.6	684 13	BJ258518 BJ258518
5	47	22.6	1013 10	BE418320 SCL024.F0
6	45.4	21.8	363 12	BG907534 TAlr1160G

C	7	45.2	21.7	468	14	BO463007	BO463007 HI02L18r
	8	45.2	21.7	479	14	BM987859	BM987859 UI-H-C00-
	9	45.2	21.7	747	13	BI914649	BI914649 603179425
	10	45.2	21.7	898	13	BI755032	BI755032 603023684
	11	44.8	21.5	988	10	BE213908	BE213908 HV_CEB000
	12	44.6	21.4	616	13	BI722737	BI722737 1031063H0
	13	44.2	21.3	686	9	AL509174	AL509174 AL509174
C	14	44.2	21.3	877	17	AO687771	AO687771 nbxb0075L
	15	43.8	21.1	568	13	BI995346	BI995346 1031026E0
	16	43.8	21.1	638	12	BG415787	BG415787 HVSMEK000
	17	43.6	21.0	421	10	BE230529	BE230529 99AS747 R
	18	43.6	21.0	985	12	BG338508	BG338508 602436229
C	19	43.4	20.9	694	13	BJ457580	BJ457580 BJ457580
	20	43.2	20.8	460	13	BM100319	BM100319 EBna01_SQ
	21	43.2	20.8	1888	11	AY104150	AY104150 Zea mays
	22	43	20.7	558	13	BM660850	BM660850 952042D06
	23	42.8	20.6	553	13	BI682538	BI682538 463649 MA
	24	42.8	20.6	563	10	BE666613	BE666613 150184 MA
	25	42.8	20.6	623	13	BM140530	BM140530 WHE0484_e
	26	42.8	20.6	648	14	BO606331	BO606331 BRY_2185
	27	42.8	20.6	788	10	BE413803	BE413803 SCU003_B1
	28	42.8	20.6	1730	13	BM479096	BM479096 AGENCOURT
C	29	42.6	20.5	925	17	CNS0091P	CNS0091P WHE2851_A
	30	42.4	20.4	448	12	BG463919	BG463919 EML_1_F11
	31	42.4	20.4	519	10	AM676807	AM676807 DGL_1_F11
	32	42.4	20.4	539	12	BG556957	BG556957 EML_40_G0
	33	42.4	20.4	589	12	BG322881	BG322881 EML_15_C0
	34	42.4	20.4	595	10	BE357434	BE357434 DGL_15_F0
	35	42.4	20.4	367	10	BE586511	BE586511 WHE0506_C
	36	42.2	20.3	367	10	BE586511	BE586511 WHE0506_C
	37	42.2	20.3	573	14	BQ246115	BQ246115 TaeL1156D
	38	42.2	20.3	630	12	BG906999	BG906999 TaeL1156D
	39	42.2	20.3	649	14	BQ246063	BQ246063 TaeL1156D
	40	42	20.2	357	14	BM080906	BM080906 1030011E0
	41	42	20.2	444	17	AG936743	AG936743 JNB-140R
	42	42	20.2	710	12	BG321153	BG321153 Zm04_03f0
	43	42	20.2	779	9	AU067540	AU067540 AU067540
	44	42	20.2	888	9	AU051760	AU051760 AU051760
	45	42	20.2	1357	11	BC029965	BC029965 Mus muscu

ALIGNMENTS

RESULT 1	AI770311/c	1282 bp	mRNA	linear	EST 24-JAN-2000
LOCUS	AI770311	42 Mycobacterium anaerobic stationary phase library	Mycobacterium smegmatis	cdna, mRNA sequence.	
DEFINITION	AI770311	42 Mycobacterium anaerobic stationary phase library	Mycobacterium smegmatis	cdna, mRNA sequence.	
ACCESSION	AI770311.1	GI:6742680			
VERSION	AI770311.1	GI:6742680			
KEYWORDS	EST.				
SOURCE	Mycobacterium smegmatis.				
ORGANISM	Mycobacterium smegmatis				
REFERENCE	1 (bases 1 to 1282)				
AUTHORS	Murugasu-Oei, B., Tay, A. and Dick, T.				
TITLE	Upregulation of stress response genes and ABC transporters in anaerobic stationary-phase Mycobacterium smegmatis				
JOURNAL	Mol. Gen. Genet. 262 (4-5), 677-682 (1999)				
COMMENT	20092472				
	Contact: Murugasu-Oei, B.				
	Mycobacterium Laboratory				
	Institute of Molecular and Cell Biology				
	30 Medical Drive, Singapore 117609, Republic of Singapore				
	Tel: 65 874 3011				
	Fax: 65 779 1117				
	Email: mcbom@imcb.nus.edu.sg				
	Insert Length: 1282 Std Error: 0.00				
	Seq primer: T3 Forward; T7 backward.				
	Location/Qualifiers				
	1. .1282				

FEATURES

source


```
/tissue_type="Leaf tissue"
/dev_stage="14 Days old"
/lab_host="E. coli XL0LR"
/note="Vector: Lambda ZapII; mass excised in plasmid
vector pBK-CMV (Stratagene).; Site_1: EcoRI; Site_2: XhoI;
mRNA obtained from wheat NIL Thatcher lrl 24 hours after
inoculation with leaf rust pathogen Puccinia triticina
race BBB carrying the avirulence gene Avr1."
BASE COUNT      162 a  178 c  179 g  130 t
ORIGIN

Query Match      22.6%; Score 47; DB 12; Length 649;
Best Local Similarity 57.0%; Pred. No. 0.97;
Matches 86; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 23 GCGCGGGTAGGCGCTACAAGGTCACAAAGACCTGGCGCTGCACGTGGCGGATCCGAT 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 188 GACGGGGTGGGCCCTACGAGCTCGGCAAGACCGCTGGCGGAGGCGAGCTTCGCGCAAGGT 247
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 83 CACCAGCTCCACGCTGACGAGGAGAGCGTGTGCGCCACCATCGAGTACCTGGTCCGCGCT 142
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 CAAGATCCCAAGGACACCCGACCGCGCCACCTGCGCCATCAAGGTGCTCGACCGCAA 307
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 143 GCACGAGGCGGCACGACGATGACCGTCCCG 173
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 CCACGCTCTCCGCCACAAAGATGGTCCGAGCAG 338
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
BJ258518
LOCUS      BJ258518      684 bp      mRNA      linear      EST 08-APR-2002
DEFINITION      aetivum cdna clone whh8d07 5', mRNA sequence.
ACCESSION      BJ258518
VERSION        BJ258518.1
KEYWORDS        EST.
SOURCE          bread wheat.
ORGANISM        Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
1 (bases 1 to 684)
Ogihara,Y. and Murai,K.
Expressed genes in Triticum aestivum
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .684
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whh8d07"
/tissue_type="spike at heading date"
/dev_stage="Feekes' scale 10.5"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown under
hydroponic conditions at UC Davis, salt stressed for 12
hours, and for 7 days, then dissected and frozen (Akhunov
in J Dvorak Lab). Total RNA was prepared from sheath
tissue, equal quantities of RNA were pooled from the two
samples, polyA was purified from the pooled RNA, a cDNA
library was made, and the cDNA clones were in vivo
excised to give pBluescript phagemids in the TJ Close lab
at the University of California, Riverside (Akhunov, Chin
Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).
Plasmid DNA preparations and DNA sequencing were

/tissue_type="Leaf tissue"
/dev_stage="14 Days old"
/lab_host="E. coli XL0LR"
/note="Vector: Lambda ZapII; mass excised in plasmid
vector pBK-CMV (Stratagene).; Site_1: EcoRI; Site_2: XhoI;
mRNA obtained from wheat NIL Thatcher lrl 24 hours after
inoculation with leaf rust pathogen Puccinia triticina
race BBB carrying the avirulence gene Avr1."
BASE COUNT      180 a  178 c  190 g  136 t
ORIGIN

Query Match      22.6%; Score 47; DB 13; Length 684;
Best Local Similarity 57.0%; Pred. No. 0.97;
Matches 86; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 23 GCGCGGGTAGGCGCTACAAGGTCACAAAGAGCTGCGCCACCATCGATCTGGTCCGCT 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 178 GACGGGGTGGGCCCTACGAGCTCGGCAAGACCGCTGGCGGAGGCGAGCTTCGCCAAGGT 237
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 83 CACCAGCTCCACGCTGACGAGGAGAGACCTGCTCGCCACCATCGATCTGGTCCGCT 142
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 238 CAAGATCGGCAAGGACACCCGCAACCGCGCCACCTGCGCCATCAAGGTGCTCGACCGCAA 297
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 143 GCACGAGGCGGCACGACGATGACCGTCCCG 173
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 298 CCACGCTCTCCGCCACAAAGATGGTCCGAGCAG 328
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
BE418320
LOCUS      BE418320      1013 bp      mRNA      linear      EST 24-JUL-2000
DEFINITION      CDNA clone SCL024.F08, mRNA sequence.
ACCESSION      BE418320
VERSION        BE418320.1
KEYWORDS        GI:9416166
SOURCE          bread wheat.
ORGANISM        Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum.
1 (bases 1 to 1013)
Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P.,
Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
Pecchioli,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
Sorreilis,M., Warburton,M. and Wenzel,G.
International Triticeae EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticeae
Unpublished (2000)
Contact: Cloutier S
Cereal Research Centre, Agriculture & Agri-Foods Canada
Winnipeg MT CANADA
Tel: 204 983 2340
Fax: 204 983 4604
Email: scloutier@em.agr.ca
International Triticeae EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.
Location/Qualifiers
1. .1013
/organism="Triticum aestivum"
/cultivar="Thatcher Lrl"
/db_xref="taxon:4565"
/clone="SCL024.F08"
/tissue_type="young leaf"
/dev_stage="after 24 hour challenge with LR-AVR1"
/note="Vector: Lambda ZAP; 1.0 Kbp average insert size."

BASE COUNT      246 a  237 c  271 g  223 t
ORIGIN

Query Match      22.6%; Score 47; DB 10; Length 1013;
Best Local Similarity 57.0%; Pred. No. 1;
Matches 86; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 23 GCGCGGGTAGGCGCTACAAGGTCACAAAGAGCTGCGCCATCGATCTGGCGGATCCGAT 82
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 184 GACGGGGTGGGCCCTACGAGCTCGGCAAGACCGCTGGCGGAGGCGAGCTTCGCCAAGGT 243
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

QY 83 CACCAGCTCCAGCTGACCGAGGAGAGCTCGTCCGCCACCATCCAGTACCTGGTCCGCT 142
DB 244 CAAGATCCGCAAGGACACCCGACCGCCACCTGCGCCATCAAGGTGCTCGACCGCAA 303

QY 143 GCACGAGGCGCCAGCAGCAGTACGCGTCCG 173
DB 304 CCACGTCTCCGCCACACAGATGGTCGAGCAG 334

RESULT 6
LOCUS BG907534 363 bp mRNA linear EST 05-JUN-2001
DEFINITION TaLr1160G09 TaLr1 Triticum aestivum cDNA clone TaLr1160G09 5',
mRNA sequence.
ACCESSION BG907534
VERSION BG907534.1 GI:14315210
KEYWORDS EST.
SOURCE bread wheat.
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 363)
AUTHORS Cloutier,S., Dong,G. and Walsh,A.
TITLE Wheat functional genomics - Thatcher Lr1 cDNA library
JOURNAL Unpublished (2001)
COMMENT Contact: Dr. Sylvie Cloutier
Cereal Research Centre, Agriculture and Agri-food Canada
195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
Tel.: (204) 983-2340
Fax: (204) 983-4604
Email: scloutier@em.agr.ca
was cloned directionally, not all sequences generated with reverse
primer were from the 5' end (same with forward primer and 3' end).
Average insert size is >2.2 kb
Plate: 160 row: G column: 09
Seq primer: M13 Reverse.
FEATURES
source
1..363
Location/Qualifiers
/organism="Triticum aestivum"
/cultivar="Thatcher Lr1"
/db_xref="taxon:4565"
/clone="TaLr1160G09"
/clone_lib="TaLr1"
/tissue_type="Leaf tissue"
/dev_stage="14 Days old"
/lab_host="E. coli XL0LR"
/notes="Vector: Lambda ZapII; mass excised in plasmid
vector pBK-CMV (Stratagene).; Site_1: EcoRI; Site_2: XhoI;
mRNA obtained from wheat NIL Thatcher Lr1 24 hours after
inoculation with leaf rust pathogen Puccinia tritici
race BBB carrying the avirulence gene Avr1."
BASE COUNT 90 a 112 c 107 g 54 t
ORIGIN
Query Match 21.8%; Score 45.4; DB 12; Length 363;
Best Local Similarity 56.3%; Pred. No. 2;
Matches 85; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 23 GCGCCGGTAGGCGGTACAAAGGTCAACAAAGCTCGCGCTGCACGTCCGCGATCCGAT 82
DB 174 GACGCGGTGGCCCTACGAGCTCGGCAAGACCTCGCGAGGCGAGCTTCTCCAAGGT 233

QY 83 CACCAGCTCCAGCTGACGAGAGAGACGTCTGTCGCCACCACGATCGAGTACCTGGTCCGCT 142
DB 234 CAAGATCGCAAGGACACCCGACCGCCACCTGCGCCATCAAGGTGCTCGACCGCAA 293

QY 143 GCACGAGGCGCCAGCAGCAGTACGCGTCCG 173
DB 294 CCACGTCTCCGCCACACAGATGGTCGAGCAG 324

RESULT 7

```

```

BQ463007 468 bp mRNA linear EST 30-MAY-2002
LOCUS HI02L18r HI Hordeum vulgare cDNA clone HI02L18 5-PRIME, mRNA
sequence.
ACCESSION BQ463007
VERSION BQ463007.1 GI:21270789
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 468)
AUTHORS Zhang,H., Weschke,W., Michalek,W., Stein,N. and Graner,A.
TITLE EST sequencing and analysis in barley (2002)
JOURNAL Unpublished (2002)
COMMENT Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 468 Std Error: 0.00
Plate: 2 row: L column: 18
Seq primer: M13rev.
FEATURES
source
1..468
Location/Qualifiers
/organism="Hordeum vulgare"
/cultivar="Barke"
/db_xref="taxon:4513"
/clone="HI02L18"
/clone_lib="HI"
/tissue_type="female inflorescences"
/dev_stage="female inflorescences (approx. 3 mm in size)"
/lab_host="XL10-Gold"
/notes="Vector: pBluescript SK+; Site_1: EcoRI (5'-end of
cDNA); Site_2: XhoI (3'-end of cDNA); Due to a cloning
artefact caused by the kit, in most cases the EcoRI site
is NOT present, as well as the EcoRI adapter used for
cloning. To excise the insert, restriction sites upstream
EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also
due to the cloning system used Blue/white selection for
recombinants is not 100% reliable."
BASE COUNT 57 a 208 c 142 g 61 t
ORIGIN
Query Match 21.7%; Score 45.2; DB 14; Length 468;
Best Local Similarity 54.1%; Pred. No. 2.3;
Matches 92; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 37 GCTACAAGTCAACAAGAAGCTCGCGCTGCACGTCCGCGATCCGATCCAGCTCCACGC 96
DB 72 GCGTCACCTCTACACACGCGGCACGCGATCGCGTTCGGCGCTTCGCGACGGCTCTCGC 131

QY 97 TGACCGAGGAAGAGCTCGTCGCCACCATCGAGTACCTGGTCCGCTGCACGAGGGCCAGC 156
DB 132 GCGCGCGCGCGACCTCCCGACACCATCTCTGCACGCGCTCCGCTCGGTACCGCACC 191

QY 157 ACACGATGACCGTCCCGGGCGCACCGAGGTGCGGCTTGAGACCGCGAC 206
DB 192 TCGACACGCGCGCATCTACGGCAGCGGCGCGCGTGGCGCGCGGCGTC 241

RESULT 8
LOCUS BM987859/c 479 bp mRNA linear EST 17-JUN-2002
DEFINITION UI-H-CO0-asx-f-09-0-UI.sl NC1_GAP_Sub9 Homo sapiens cDNA clone
IMAGE:5860361 3', mRNA sequence.
ACCESSION BM987859
VERSION BM987859.1 GI:19707248
KEYWORDS EST.
SOURCE human.

```

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 479)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov Tissue procurement: Dr. Jose Mercuende CDNA Library preparation: Dr. M. Bento Soares, University of Iowa CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Clone Distribution Information can be found through the I.M.A.G.E. Consortium/LINL at: http://image.llnl.gov The following repetitive elements were found in this cDNA sequence: 6-67, >MLTCL#LTR/MaLR Seq primer: M13 FORWARD POLYA=Yes.

FEATURES	source	Location/Qualifiers	1. .479	80 a	155 c	111 t	1 others
		/organism="Homo sapiens"					
		/db_xref="taxon:9606"					
		/clone="IMAGE:5860361"					
		/clone_lib="NCI_CGAP_Sub9"					
		/tissue_type="mixed"					
		/dev_stage="mixed"					
		/lab_host="DH10B (Life Technologies)"					
		/note="Vector: pT7T3-Pac (Pharmacia) with a modified polylinker. Site.1: EcoR I; Site.2: Not I; tissues: Cholonic mucosa with Crohns disease, Cholonic mucosa with ulcerative colitis, Fetal thymus, Cervix, Cervical adenosquamous carcinoma, Ligament cells, Prostate carcinoma, Bladder carcinoma, Brain oligodendroga ; NCI_CGAP_Sub9 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are CGTC, AACG, GGCGCC, GGAAG, TAGC, TAAGC, ATGCG, AGACA, ATCAC. For additional information, contact: Bento Soares, bento-soares@uiowa.edu					
		TAG_LIIB-UI-H-CO0					
		TAG_TISSUE=Bladder Carcinoma					
		TAG_SEQ=AGACA*					

[illegible]

RESULT 9			
BI914649			
LOCUS	BI914649	747 bp	mRNA linear EST 16-OCT-2001

DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

603179425F1 NTH_MGC_121 Homo sapiens cDNA clone IMAGE:5243431 5',
mRNA sequence.
BI914649
BI914649.1 GI:16178742
EST.
human.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
ORGANISM				

Homo sapiens
Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 747)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11614 row: b column: 08
High quality sequence stop: 747.

FEATURES	SOURCE
-----------------	---------------

```

1. .747
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5243431"
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
/note="organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 017. Not
this is a NIH_MGC Library."
177 a 194 c 275 g 101 t

```

BASE COUNT
ORIGIN

177 a 194 c 275 g 101 t

21.7%; Score 45.2; DB 13; Length 747;
Similarity 54.9%; Pred. No. 2.4;

Matches

89: Conservative 0; Mismatches 73; Indels 0; Gaps

Qy 75 0

TCCGATCACAGGTCCACCGTGCACGGAGGACGTCGCCACCATCGAGTACCTG 134
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
AGGCCATCGCGCGCTCATGCCAAGTCGTGCGGTCCTGGAGATCCTGAGGTGCTG 184

Qy 135 0

CCCGCCTGACGAGGGCCAGCACAGATGACCGTCCCGGC 176
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 CAGCGGCCACACGTCGCGCCGAGCTGGACGAGCTCGC 226

1
2
3
4
5
6

BI755032 898 bp mRNA linear EST 25-SEP-2001
603023684f1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5194052 5',
mRNA sequence.

ACCESSION
VERSION

BI755032
BI75032.1 GI:15746610
EST.
human.

[illegible]

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.


```

ACCESSION   BI722737
VERSION     BI722737.1  GI:15698432
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
ORGANISM    Chlamydomonas reinhardtii.
REFERENCE   Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
AUTHORS     P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
TITLE       Analyses of the Chlamydomonas reinhardtii Genome: A Model,
            Unicellular System for Analyzing Gene Function and Regulation in
            Vascular Plants. Project: 1031
JOURNAL     Unpublished (2001)
COMMENT     Contact: Charles Hauser
            DCMB Box 91000
            Durham, NC 27708-1000
            Tel: 919 613 8159
            Fax: 919 613 8177
            Email: chauser@duke.edu.
FEATURES    source
            Location/Qualifiers
            1..616
            /organism="Chlamydomonas reinhardtii"
            /strain="CC-1690 wild type mt+ 21gr"
            /db_xref="taxon:3055"
            /clone_lib="C. reinhardtii CC-1690, Stress II (normalized
            ), Lambda Zap II"
            /note="Vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
            XhoI; Stress condition II library, constructed by John
            Davies and Jeffrey McDermott, combines cDNAs from CC-1690
            cells grown to mid-log phase in TAP (NH4+ - containing)
            and shifted to TAP - NO3- (24hrs); H2 production
            conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
            Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
            sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
            POLYA mRNA was purified from each sample, pooled and cDNA
            synthesized. The cDNA was directionally cloned into lambda
            Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
            sites. pBluescript II SK- plasmids were excised from the
            lambda Zap clones by superinfection with ExAssist
            (Stratagene) phage. The library was normalized using
            method 4 described in Bonaldo et al., (1996) Genome
            Research 6: 791-806."
            Research 6: 791-806."
BASE COUNT  156 a 170 c 207 g 83 t
ORIGIN
Query Match      21.4%; Score 44.6; DB 13; Length 616;
Best Local Similarity 51.2%; Pred. No. 3.2;
Matches 104; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 1 TCAGGAGAGGGCTACGACCTGGCCCGGTAGGCGCGGTACAAGGTCAACAAGAGCTCG 60
    || ||||| || || || || || || || || || || || || || || || || ||
Db 138 TCTGGAGACAGTTTGGCCGCGCCATCAAGCTGGGCATCATCGAGGACCAACCAACGCA 197
    || || || || || || || || || || || || || || || || || || || || ||
QY 61 GCCTCAGCTGGCGGATCCGATCACCAGCTCCAGCTGACCGAGGAGACAGCTGCTGCCA 120
    || || || || || || || || || || || || || || || || || || || || ||
Db 198 ACCGCTGCCAAGTTGCTGGCTTCCACACCTCCAGACCGCGGACCAAGCTCACCACCC 257
    || || || || || || || || || || || || || || || || || || || || ||
QY 121 CCATCAGTACCTGTGTCGCCCTGCACGAGGCGCCAGCACGATGACCGTCCCGGGCGGCA 180
    || || || || || || || || || || || || || || || || || || || || ||
Db 258 TGGACGAGTACATCGGCCGTCATGAAGAGGCGCCAGAGTCCATCTACTACTGCGCGGCA 317
    || || || || || || || || || || || || || || || || || || || || ||
QY 181 CCGAGGTGCGGTTGAGACCCAC 203
    || || || || || || || || || || || || || || || || || || || || ||
Db 318 CCAGCAAGGAGGAGGTGGCTGCTGC 340
    || || || || || || || || || || || || || || || || || || || || ||

RESULT 13
AL509174
LOCUS
DEFINITION   AL509174 Hordeum vulgare Barke developing caryopsis (3.-15.DAP)
            Hordeum vulgare cDNA clone HV10020V 5', mRNA sequence.
            686 bp mRNA linear EST 04-JAN-2001
AUTHORS      AL509174
            AL509174 Hordeum vulgare Barke developing caryopsis (3.-15.DAP)
            Hordeum vulgare cDNA clone HV10020V 5', mRNA sequence.

```

```

AL509174
AL509174.1  GI:12035677
EST.
Hordeum vulgare.
Hordeum vulgare.
ORGANISM    Hordeum vulgare.
REFERENCE   1 (bases 1 to 686)
AUTHORS     Michalek,W., Weschke,W., Pleissner,K.-P. and Graner,A.
TITLE       EST sequencing and analysis in barley
JOURNAL     Unpublished (2000)
COMMENT     Contact: Michalek W
            Institute for Plant Genetics and Crop Plant Research
            Corrensstr.3, D-06466 Gatersleben, Germany
            Email: michalek@ipk-gatersleben.de, http://pgrc.ipk-gatersleben.de
            Seq primer: T3 primer for 5'end.
FEATURES    source
            Location/Qualifiers
            1..686
            /organism="Hordeum vulgare"
            /cultivar="Barke"
            /db_xref="taxon:4513"
            /clone_lib="Hordeum vulgare Barke developing caryopsis
            (3.-15.DAP)"
            /tissue_type="developing caryopsis (3.-15.DAP)"
            /lab_host="XLOLR"
            /note="Vector: plasmid pBK-CMV; Site_1: EcoRI; Site_2:
            XhoI; mRNA was made from developing caryopsis (3.-15.DAP)
            of spring barley variety 'Barke', a high quality malting
            variety. Cloning sites: EcoRI (5'-end of cDNA) and XhoI
            (3'-end of cDNA). NOTE: Due to a cloning artefact caused
            by the kit, in most cases the EcoRI site is NOT present,
            as well as the EcoRI adapter. Average insert size is 1 kb
            Sequence trimming: vector sequences and sequence ends were
            trimmed from the 5'- and 3'-end until a 50 bp window
            contains less than two ambiguities. The maximum length was
            set to 700 bp"
BASE COUNT  92 a 265 c 223 g 98 t 8 others
ORIGIN
Query Match      21.3%; Score 44.2; DB 9; Length 686;
Best Local Similarity 53.5%; Pred. No. 4;
Matches 91; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 37 GCTACAAGGTCAACAAGAGCTCGGCTGCACGTCGGCATCGATCCAGCTCCACGC 96
    || || ||||| || || || || || || || || || || || || || || || || ||
Db 80 GCGTCAACCTCAACACGGGCCACGCGATGCCGNGCTGGCTTCGGCACCGGCTCTCGC 139
    || || || || || || || || || || || || || || || || || || || || ||
QY 97 TGACCGAGGAAGACGTGCTGTCGCCACCATCGAGTACCTGTCGCCCTGCACGAGGCCACG 156
    || || || || || || || || || || || || || || || || || || || || ||
Db 140 GCGCGCGCGCGACCTGCCGACACCATCTGCACGCGCTCCGCTCGGTACCGCCACC 199
    || || || || || || || || || || || || || || || || || || || || ||
QY 157 ACAGCATGACCGTCCCGGGCGGACCGAGGTGCCGTTGAGACCCGACGAC 206
    || || || || || || || || || || || || || || || || || || || || ||
Db 200 TCGACAGCGCGCCATGTACGGCACGGAGCCGCGCGCGCGCGTC 249
    || || || || || || || || || || || || || || || || || || || || ||

RESULT 14
AL509174
LOCUS
DEFINITION   AL509174 Hordeum vulgare Barke developing caryopsis (3.-15.DAP)
            Hordeum vulgare cDNA clone HV10020V 5', mRNA sequence.
            877 bp DNA linear GSS 01-JUL-1999
AUTHORS      AL509174
            AL509174 Hordeum vulgare Barke developing caryopsis (3.-15.DAP)
            Hordeum vulgare cDNA clone HV10020V 5', mRNA sequence.

```

TITLE
JOURNAL
COMMENT

A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCATATAGG
Class: BAC ends
High quality sequence stop: 85.
Location/Qualifiers
source
1. .877
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="nbx0075L01f"
/clone_lib="CUGI Rice BAC Library"
/tissue_type="Leaf"
/lab_host="E. coli DH10B"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."
BASE COUNT 107 a 370 c 257 g 143 t
ORIGIN
Query Match 21.3%; Score 44.2; DB 17; Length 877;
Best Local Similarity 53.8%; Pred. No. 4.1;
Matches 91; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
Qy 33 GGCGCTACAAAGTCAACAAAGACTCGCGCTGCAGCTCGCGCATCCGATCACCAGCTCC 92
Db 232 GCGCGGAGGCGACCAACCCAGGCGGTGTGACCGTGTGCTCGGAGTTCTCTCGGACGTTTC 173
Qy 93 ACCTGACCGAGGAAGACGTCTGTCGCCACCATCGAGTACCTGTCTCGCTGACAGAGGCG 152
Db 172 CAGGGATGGAGCGGAGCTGGAGGCGGACATGGAGCGCATGGGCGGCCCGGACGAGCGC 113
Qy 153 CAGCACAGTACCGCTCCCGGGCGGACCGAGGTGCGCGTTTGAGACCG 201
Db 112 CTGCTCCGGCGACGACGCGCGCTGGGCGCGAGCTTCCGAGGCGG 64
RESULT 15
BI995346
LOCUS
DEFINITION
1031026E07.y2 C. reinhardtii CC-1690, Stress II (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BI995346 568 bp mRNA linear EST 25-OCT-2001
1031026E07.y2 C. reinhardtii CC-1690, Stress II (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
GI:16430136
Chlamydomonas reinhardtii.
Chlamydomonas reinhardtii.
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

1. (bases 1 to 568)
Grossman,A., Chang,C.-W., Davies,J., Harris,E., Hauser,C., Lefebvre
P., McDermott,J.P., Shrager,J., Silflow,C. and Stern,D.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1031
Unpublished (2001)
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.
Location/Qualifiers
1. .568
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Stress II (normalized
) , lambda Zap II"
/note="vector: pBluescript II SK-; Site_1: EcoRI; Site_2:
XhoI; Stress condition II library, constructed by John
Davies and Jeffrey McDermott, combines cDNAs from CC-1690
cells grown to mid-log phase in TAP (NH4+ - containing)
and shifted to TAP - NO3- (24hrs); H2 production
conditions (0, 12hr, 24hr) see Melis et al., (2000) Plant
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).
POLYA mRNA was purified from each sample, pooled and cDNA
synthesized. The cDNA was directionally cloned into lambda
Zap II (Stratagene) in the EcoRI (5') and XhoI (3')
sites. pBluescript II SK- plasmids were excised from the
lambda Zap clones by superinfection with ExAssist
(Stratagene) phage. The library was normalized using
method 4 described in Bonaldo et al., (1996) Genome
Research 6: 791-806."
BASE COUNT 80 a 206 c 185 g 97 t
ORIGIN
Query Match 21.1%; Score 43.8; DB 13; Length 568;
Best Local Similarity 50.7%; Pred. No. 4.7;
Matches 105; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
Qy 2 CAAGGAGAGCGCTACGACCTGGCCCGGTAGCGCGCTACAAGGTCAACAGACCTCGG 61
Db 251 CAAGGAGAGCGCTTCGAGGTGTGTCACGCGCTGTGAACAGGCTCATGGAGGTCTCGG 310
Qy 62 CCGTCCAGCTCGCGCATCCGATCCAGCTCCAGCTGACCGAGGAGAGACGTCTGTCGCCAC 121
Db 311 CGTCCGTACAGGAGAGCGCGCGCTCTCCCGCGCGCTGCTCTACCACTGGCGCC 370
Qy 122 CATCGAGTACCTGTGTCGCCCTGACGAGGGCGGACACAGATGACCTGTCGCCGGGGGCG 181
Db 371 CGCCGCCGACCCCGCGCTGTTCGCCGGCGGCGCGCGGTGTTCCGCTCCACTCGCGG 430
Qy 182 CGAGTGCCTGGTTGAGACCGAGCAT 208
Db 431 CGCAAGCCCGTGGCGGTGTTCCGGCAT 457
Search completed: November 13, 2002, 03:59:26
Job time : 1134.63 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 ; Search time 21.1607 Seconds
(without alignments)
3487.380 Million cell updates/sec

Title: US-09-697-123B-1

Perfect score: 208

Sequence: 1 tcaagagaagcgtacgac.....coggttgagaccgacacat 208

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications_NA:*
- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
 - 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/2/pubpna/US10_PUB.seq:*
 - 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
 - 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
 - 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	83.2	40.0	5096	10	US-09-984-711-5		Sequence 5, Appli
2	83.2	40.0	5099	9	US-10-075-460-5		Sequence 5, Appli
3	83.2	40.0	5099	10	US-09-887-052-1		Sequence 1, Appli
4	83.2	40.0	5099	10	US-09-887-052-3		Sequence 3, Appli
5	83.2	40.0	5099	10	US-09-887-052-5		Sequence 5, Appli
6	44.4	21.3	2329	10	US-09-816-828-9		Sequence 9, Appli
7	38.4	18.5	1998	10	US-09-815-242-7695		Sequence 1, Appli
8	37.2	17.9	1066	10	US-09-879-833-1		Sequence 1, Appli
9	37.2	17.9	1437	10	US-09-815-242-7738		Sequence 7738, Ap
10	37.2	17.9	2451	10	US-09-771-1614-4		Sequence 4, Appli
11	37.2	17.9	4444	10	US-09-879-833-3		Sequence 3, Appli
12	36.6	17.6	526	10	US-09-731-872-190		Sequence 190, App
13	36.6	17.6	654	10	US-09-731-872-229		Sequence 229, App
14	36.4	17.5	1314	10	US-09-815-242-7740		Sequence 7740, Ap
15	36.2	17.4	427	10	US-09-960-352-5073		Sequence 5073, Ap
16	36.2	17.4	978	10	US-09-815-242-8005		Sequence 8005, Ap
17	36.2	17.4	1266	10	US-09-815-242-7920		Sequence 7920, Ap
18	36.2	17.4	3731	10	US-09-832-268A-1		Sequence 1, Appli
19	36	17.3	1004	10	US-09-755-830-7		Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-09-984-711-5
; Sequence 5, Application US/09984711
; Patent No. US20020119549A1

; GENERAL INFORMATION:

; APPLICANT: MOECKEL, Bettina

; APPLICANT: BATHE, Brigitte

; APPLICANT: STEPHAN, Hans

; APPLICANT: KREUTZER, Caroline

; APPLICANT: HERMANN, Thomas

; APPLICANT: PFEFFERLE, Walter

; APPLICANT: BINDER, Michael

; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE

; FILE REFERENCE: 204209US0

; CURRENT APPLICATION NUMBER: US/09/984,711

; CURRENT FILING DATE: 2001-10-31

; PRIOR APPLICATION NUMBER: DE10108230.9

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 5

; LENGTH: 5096

; TYPE: DNA

; ORGANISM: Corynebacterium glutamicum

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (702)..(4196)

; OTHER INFORMATION:

US-09-984-711-5

Query Match 40.0% Score 83.2; DB 10; Length 5096;

Best Local Similarity 67.0%; Pred.No. 1.8e-12;

Matches 134; Conservativ: 0; Mismatches 63; Indels 3; Gaps 1;

Qy 9 AAGCGCTACGACCTGGCCGGGTAGCGGTCTACAAAGTCAACAAAGCTGGCGCTGCAC 68

|||||

Db 1578 AAGCGCTACGACCTGGCGGTGGTGTGTACAAAGTCAACAAAGCTGGCGCTT--- 1634

|||||

Qy 69 GTCGGCGATCGATCACCAGCTCCACGCTGACCGAGGAGGAGCTGCGCGCCACCATCGAG 128

|||||

Db 1635 GGTGGCGACCATGATGTTGATGACTCTTACTGAAGGAGGACATCGCAACCATCGAG 1694

|||||

Db 1635 GGTGGCCACACGATGTTTGTATGACTCTTACTGAAGAGGACATCGCAACCCACCATCGAG 1694
Qy 129 TACCTGGTCCGCTGACGAGGCCAGCACACGATGACCGTCCCGGGCGGCACCGGAGTG 188
Db 1695 TACCTGGTGGCTGTCACGCGAGGTGAGCGGCTCATGACTTCTCCAAATGGTGAAGAGATC 1754
Qy 189 CCGGTTGAGACCGGACGACAT 208
Db 1755 CCAGTCGAGCCGATGACAT 1774
RESULT 5
US-09-887-052-5
; Sequence 5, Application US/09887052
; Patent No. US20020119537A1
; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina
; APPLICANT: BATHE, Brigitte
; APPLICANT: HERMANN, Thomas
; APPLICANT: PEFFERLE, Walter
; APPLICANT: BINDER, Michael
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rPOB GENE
; FILE REFERENCE: 204212U50X
; CURRENT APPLICATION NUMBER: US/09/887,052
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: DE10107229.5
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (702)..(4196)
US-09-887-052-5
Query Match 40.0%; Score 83.2; DB 10; Length 5099;
Best Local Similarity 67.0%; Pred. No. 1.8e-12;
Matches 134; Conservative 0; Mismatches 63; Indels 3; Gaps 1;
Qy 9 AAGCGCTACGACCTGGCCGGGTAGCGCTACAAGGTCAACAAGAGCTCGGCGCTGCAC 68
Db 1578 AAGCGCTACGACCTGGCTCGGTTGGTGTACAAAGATCAACCGCAGCTCGGCGCTT--- 1634
Qy 69 GTCGGCGATCCGATCACACGCTCCAGCTGACCGAGGAGACGCTGCTGCGCCACCATCGAG 128
Db 1635 GGTGGCAGCCAGATGTTTGTATGACTCTTACTGAAGAGGACATCGCAACCCACCATCGAG 1694
Qy 129 TACCTGGTCCGCTGACGAGGGCCAGCACACGATGACCGTCCCGGGCGGCACCGGAGTG 188
Db 1695 TACCTGGTGGCTGTCACGCGAGGTGAGCGGCTCATGACTTCTCCAAATGGTGAAGAGATC 1754
Qy 189 CCGGTTGAGACCGGACGACAT 208
Db 1755 CCAGTCGAGCCGATGACAT 1774
RESULT 6
US-09-816-828-9
; Sequence 9, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyang
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Zhiwei

; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 791CIP2E
; CURRENT APPLICATION NUMBER: US/09/816,828
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 9
; LENGTH: 2329
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1612)..(2211)
US-09-816-828-9
Query Match 21.3%; Score 44.4; DB 10; Length 2329;
Best Local Similarity 55.1%; Pred. No. 0.0066;
Matches 87; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
Qy 51 AAGAAGCTCGGCTGCACGTCGGCGATCGATCACGATCCACGCTCCACGAGGAAGAC 110
Db 1924 AAGGCCAGCGCAAGAAGGTGGCCGACGCGCTGACCAACGCGCTGGCGCACGTCGACGAC 1983
Qy 111 GTCTGCGCCACCATCAGTACCTGTGTCGCGCTGTCGAGGGCCAGCACACGATGACCGTTC 170
Db 1984 ATGCCCAACGCGTGGCGCACGTGGACACATGCCCAACGCGCTGCCGCCCTGAGCGAC 2043
Qy 171 CCGGGCGCAGCGAGTGGCGGTTGAGACCGGACGACAT 208
Db 2044 CTGCAGCGCACAAAGCTTCGGGTGGACCGGTCAACTT 2081
RESULT 7
US-09-815-242-7695
; Sequence 7695, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.O11A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7695
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1998)
US-09-815-242-7695

Query Match 18.5%; Score 38.4; DB 10; Length 1998;
Best Local Similarity 51.8%; Pred. No. 0.2;
Matches 87; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
Qy 17 CGACTGCGCGGTAGCGCTACAAGGTCAACAAGAAGCTGGCGCTGCACGTGCGCGA 76
Db 1683 CGGCTGGCGCTGAGCGCTACGACAAGCTCAGGACAGCGGCGCAAGTCCGGTGGT 1742
Qy 77 TCCGATCACCGTCCACGCTGACCGAGGAGAGCTGCTGCCACCATCGAGTACCTGGT 136
Db 1743 ATCGATGCCATGCCACCGCTTACGAGCAGCAGGAGTCTTACAAAGCAGTCCGCTGCT 1802
Qy 137 CCGCTGCACGAGCGCCAGCAGCATGACCGTCCCGGCGGCACCA 184
Db 1803 GCGGTGAAGTCGCGCGCGCATCGGCATCGAGGCGCGCCATGCCGA 1850

RESULT 8
US-09-879-833-1
; Sequence 1, Application US/09879833
; Patent No. US20020055107A1
; GENERAL INFORMATION:
; APPLICANT: Park, Chong-Hun; SAMYANG GENEX CORPORATION
; TITLE OF INVENTION: CANCER DIAGNOSTIC METHOD USING CELL GROWTH INHIBITING AND CELL DEATH INDUCING AGENTS
; TITLE OF INVENTION: FERMENTATION SPECIFIC SYG972 GENE AND GENOMIC DNA AND PROMOTER THEREOF
; TYPE: DNA
; FILE REFERENCE: PA/SYG/99601
; CURRENT APPLICATION NUMBER: US/09/879,833
; PRIOR FILING DATE: 2001-06-12
; PRIOR FILING DATE: KR 54933
; PRIOR FILING DATE: 1998-12-12
; PRIOR FILING DATE: KR 63958
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: KOPATIN 1.5
; SEQ ID NO 1
; LENGTH: 1066
; TYPE: DNA
; ORGANISM: HL60 cell line
US-09-879-833-1

Query Match 17.9%; Score 37.2; DB 10; Length 1066;
Best Local Similarity 52.6%; Pred. No. 0.38;
Matches 81; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
Qy 55 AGCTCGGCTGCACGTGCGCGATCGGATCCAGCTCCAGCTGACCGAGGAGACGCTCG 114
Db 294 AGGAGGAGGAGGCGACATCGCGCTGCGATCCATTTTACGCTGATCCAGGCTTTCTGCT 353
Qy 115 TCGCCACCATCGAGTACCTGCTGCGCTGACGAGGCGCCACACGATGACCGTCCCGG 174
Db 354 GCGAAGACGATCGACATATGTCGCGTGGCGATGTGACGCGCTGGCGGCTATCTGTTG 413
Qy 175 GCGGACCGAGTGTCGCGTGTGAGACCGGACGACAT 208
Db 414 GCGCGGAGGAGGCGGCTGCGCGGCGACCT 447

RESULT 9
US-09-815-242-7738
; Sequence 7738, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in Pseudomonas
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7738
; LENGTH: 1437
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1437)
US-09-815-242-7738

Query Match 17.9%; Score 37.2; DB 10; Length 1437;
Best Local Similarity 54.3%; Pred. No. 0.39;
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
Qy 46 TCAACAAGAAGCTGCGCGCTGCGAGTCCGCGATCCAGCTCCAGCTGACCGAGG 105
Db 809 TCGACAAGCTGATGCTGCGCGTCCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 868
Qy 106 AAGACGTGCTGCGCCACCATCGAGTACCTGCTGCGCTGCGCGCTGCGCGCTGCGCG 165
Db 869 ACAGCGCGTGAGCCCTGCGAGCGCTGCTTATCTACGTGCGACGACCATCGACGATGA 928
Qy 166 CCGTCCCGGCGGCGACCG 183
Db 929 GCGTTCGCGGCGCTTCG 946
RESULT 10
US-09-771-161A-4
; Sequence 4, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 2451
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: -
; LOCATION: (1)..(2451)
; OTHER INFORMATION: "n" can be any nucleotide 'a', 'c', 'g' or 't'
US-09-771-161A-4

Query Match          17.9%; Score 37.2; DB 10; Length 2451;
Best Local Similarity 53.4%; Pred. No. 0.4;
Matches 78; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 28 GGGTAGGCGCTACAGGCTCAACAAGAGCTCGCGCTGCAGCTCGCGCATCCGATCACCA 87
    || || || || || || || || || || || || || || || || || || || || ||
Db 1748 GGATGAGCAACCCCGAGGTCTCCGCAACCTGGAGCGGGCTACCGCATGCGCGGCCCG 1807
    || || || || || || || || || || || || || || || || || || || || ||
QY 88 GTCTCACGCTGACCCAGGAGAGAGCTGCTGCGCCACCATCGAGTACCTGGTCCGCGCTGCACG 147
    || || || || || || || || || || || || || || || || || || || || ||
Db 1808 ACACCTGCGCGCGCGAGCTGTACCGCGGGGTCTATCGCGAGTGTGCGCGACCGCGCGCG 1867
    || || || || || || || || || || || || || || || || || || || || ||
QY 148 AGGGCAGCACACGATGACCGTCCCG 173
    || || || || || || || || || || || || || || || || || || || || ||
Db 1868 AGGAGCGGCCACCTTCGAGTTCTCTG 1893
    || || || || || || || || || || || || || || || || || || || || ||

RESULT 11
US-09-879-833-3
; Sequence 3, Application US/09879833
; Patent No. US2002055107A1
; GENERAL INFORMATION:
; APPLICANT: Park, Chong-Hun; SAMYANG GENEX CORPORATION
; TITLE OF INVENTION: CANCER DIAGNOSIS METHOD USING CELL GROWTH INHIBITING AND CELL DIF
; FILE OF INVENTION: FERTILIZATION SPECIFIC SYG972 GENE AND GENOMIC DNA AND PROMOTER T
; TITLE OF INVENTION: EROF
; FILE REFERENCE: PA/SYG/99601
; CURRENT FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: US/09/879,833
; PRIOR FILING DATE: 1998-12-12
; PRIOR APPLICATION NUMBER: KR 54933
; PRIOR FILING DATE: 1998-12-12
; PRIOR APPLICATION NUMBER: KR 63958
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: KOPATIN 1.5
; SEQ ID NO 3
; LENGTH: 4444
; TYPE: DNA
; ORGANISM: HL60 cell line
US-09-879-833-3

Query Match          17.9%; Score 37.2; DB 10; Length 4444;
Best Local Similarity 52.6%; Pred. No. 0.42;
Matches 81; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 55 AGCTGGGCTGCACGTGGCGATCCGATCACCAGCTCCAGCTGACCCGAGGAGACGTCG 114
    || || || || || || || || || || || || || || || || || || || || ||
Db 3669 AGAGGACGAGGCGACATCGCGCTGCAGATCCATTTTACGCTGATCCAGGCTTTCTGCT 3728
    || || || || || || || || || || || || || || || || || || || || ||
QY 115 TCGCCACCATCGAGTACCTGGTCCGCTGCAGAGGGCCGACGACGATGACCGGTCCCGG 174
    || || || || || || || || || || || || || || || || || || || || ||
Db 3729 GCGAGAACACATGACATAGTGCCTGGCGGTGGGATGTGCAGCGGGCTGCGGCTATCGTGG 3788
    || || || || || || || || || || || || || || || || || || || || ||
QY 175 GCGGACCGAGTGGCGGTGTGAGACCGGAGGACAT 208
    || || || || || || || || || || || || || || || || || || || || ||
Db 3789 GCGCGCGGAGGAGGCGGTGCGCGGCGGCGACCT 3822
    || || || || || || || || || || || || || || || || || || || || ||

RESULT 12
US-09-731-872-190
; Sequence 190, Application US/09731872
; Patent No. US20020102604A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 190
; LENGTH: 526
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..368
US-09-731-872-190

Query Match          17.6%; Score 36.6; DB 10; Length 526;
Best Local Similarity 54.0%; Pred. No. 0.51;
Matches 75; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 22 TGGCCCGGGTAGGCGCTACAGGCTCAACAAGAGCTCGCGCTGCAGCTCGCGCATCCGA 81
    || || || || || || || || || || || || || || || || || || || || ||
Db 64 TGGCTGGGGCGCGCTACTGCTGTACGACGAGAGCTGCTGGGGCCCGCAGCACAAGA 123
    || || || || || || || || || || || || || || || || || || || || ||
QY 82 TCACCAAGCTCCAGCTGACCGAGGAGGAGGAGCTGCTGCGCCACCATCGAGTACCTGGTCCGCC 141
    || || || || || || || || || || || || || || || || || || || || ||
Db 124 GCCAGGAGCGCTTACAGAGGCTGGGGAGGTGGTCCCGCCCGCATGTACCAAGTTACGCC 183
    || || || || || || || || || || || || || || || || || || || || ||
QY 142 TGCACGAGGCGCCAGCACAC 160
    || || || || || || || || || || || || || || || || || || || || ||
Db 184 AGTACGTGTGTACGACAGAC 202
    || || || || || || || || || || || || || || || || || || || || ||

RESULT 13
US-09-731-872-229
; Sequence 229, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 229
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 86..505
US-09-731-872-229

Query Match          17.6%; Score 36.6; DB 10; Length 654;
Best Local Similarity 54.0%; Pred. No. 0.52;
Matches 75; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 22 TGGCCCGGGTAGGCGCTACAGGCTCAACAAGAGCTCGCGCTGCAGCTCGCGCATCCGA 81
    || || || || || || || || || || || || || || || || || || || || ||
```


Db 201 TGGCTGGGGCCCGCTTACCTGGTGTACGACCCAGGAGCTGCTGGGGCCCGCCAGACAAGA 260
QY 82 TCACCAAGCTCCACGCTGACCCGAGGAAGAGCGTGTGCGCCACCATCGAGTACCTGGTCCGCC 141
Db 261 GCCAGGCACCCCTACAGAGGCTGGGAGGTGGTCCCGCCCGCCATGTACCAAGTTCAGCC 320

QY 142 TCACAGCGGGCCAGCACAC 160
Db 321 AGTACGTGTGTGAGCAGAC 339

RESULT 14

US-09-815-242-7740

; Sequence 7740, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US-09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows version 4.0

; SEQ ID NO 7740

; LENGTH: 1314

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)...(1314)

; US-09-815-242-7740

Query Match 17.5%; Score 36.4; DB 10; Length 1314;
Best Local Similarity 51.9%; Pred. No. 0.61;
Matches 82; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 26 CCGGGTAGCGCGCTACAGGTCAACAGAGCTCGGCGCTGACGTGCGCGGATCCGATCAC 85
Db 612 CGGTATCGTATCGCCACCGCCATTCGCGACCGCGGCATCAGCGACGCGGCTGGTGCAC 671

QY 86 CAGCTCCACCGCTACCGAGGAAGACGTGTCGCCACCATCGAGTACCTGTCGCGCTGCA 145
Db 672 GCGCTATTCATGATCGCCGGGTGTCGCGCCACCGCGGCTGTCCCTGTCTACCTGGC 731

QY 146 CGAGGGCCAGCACAGATGACCGTCCCGGGCGGCACCG 183
Db 732 GCTGTTCTATCTTGGCGCCACAGCCAGGGGCATCGCG 769

RESULT 15

US-09-960-352-5073

; Sequence 5073, Application US/09960352

; Patent No. US20020137139A1

; GENERAL INFORMATION:

; APPLICANT: Warren, Wesley C.

; APPLICANT: Tao, Nengbing

; APPLICANT: Byatt, John C.

; APPLICANT: Mathialagan, Nagappan

; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION

; FILE REFERENCE: 16511.006/37-21(10298)C

; CURRENT APPLICATION NUMBER: US/09/960,352

; CURRENT FILING DATE: 2001-09-24

; NUMBER OF SEQ ID NOS: 15112

; SEQ ID NO 5073

; LENGTH: 427

; TYPE: DNA

; ORGANISM: Bos taurus

; OTHER INFORMATION: Clone ID: 22-LIB3058-020-Q1-K1-F9

; US-09-960-352-5073

Query Match 17.4%; Score 36.2; DB 10; Length 427;
Best Local Similarity 49.2%; Pred. No. 0.63;
Matches 95; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 8 GAAGCGCTACGACCTGGCCCGGTAGCGCTACAGGTCAACAGAGCTCGCGCTGCA 67

Db 26 GCACAGCGCGGCTTCAACCTGGAGGCTGCCGCGAGTACAGCGCTCATCGTTACAA 85

QY 68 CGTCGGCGATCCGATCACCAGCTCCAGCTGACCGAGAGAGAGTCTGTCGCCACCATCGA 127

Db 86 CGCCATCGACTCGTGTGACGCGCATCTGTCGCGCCCTGCGCGCTCAAGATCGACTTCCA 145

QY 128 GTACCTGTGCGCTGTGACGAGGCCAGACAGATGACCGTCCCGGGCGGCACCGAGGT 187

Db 146 CAACCCCGACCGCGCTACGACCGCTGTCGCGCTACGCGGCGCCGCGAGAG 205

QY 188 GCCGTTGAGACC 200

Db 206 CAAGGGCGAGATC 218

Search completed: November 12, 2002, 16:57:36

Job time : 35.1607 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:56:36 : Search time 27.8936 Seconds
(without alignments)
2886.857 Million cell updates/sec

Title: US-09-697-123B-1
Perfect score: 208
Sequence: 1 tcaaggagagcgctacgac.....ccggttagagaccgacgacat 208

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 1533381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	166.4	80.0	970	1	US-08-250-030-1
2	166.4	80.0	970	5	PCT-US95-06790-1
3	142.4	68.5	3447	2	US-08-313-185-57
4	142.4	68.5	3447	3	US-09-082-614A-57
5	49.8	23.9	6085	4	US-09-029-603-4
6	45.4	21.8	4403765	4	US-09-103-840A-2
7	43	20.7	2122	4	US-09-029-603-1
8	42	20.2	2244	1	US-08-476-519-10
9	42	20.2	2244	5	PCT-US95-09323-10
10	42	20.2	2334	1	US-08-476-519-1
11	42	20.2	2334	5	PCT-US95-09323-1
12	40.2	19.3	4403765	4	US-09-103-840A-2
13	40.2	19.3	4411529	4	US-09-103-840A-1
14	39.2	18.8	12588	2	US-08-387-942C-1
15	39	18.8	44377	2	US-08-804-227C-7
16	39	18.8	44377	2	US-08-804-198-1
17	38.4	18.5	1998	4	US-09-382-106-1
18	38	18.3	765	3	US-08-718-904-79
19	37.4	18.0	20235	1	US-07-642-734C-3
20	37.4	18.0	20235	3	US-08-439-009A-3
21	37.2	17.9	477	1	US-08-797-831A-2
22	37.2	17.9	477	2	US-08-463-081B-31
23	37.2	17.9	477	2	US-08-461-379A-31
24	37.2	17.9	477	2	US-08-462-390B-31
25	37.2	17.9	477	3	US-08-463-074B-31
26	37.2	17.9	477	3	US-08-465-585C-31
27	37.2	17.9	477	3	US-08-652-446-31

28	37.2	17.9	1036	1	US-08-797-831A-3	Sequence 3, Appli
29	37.2	17.9	1065	2	US-08-463-081B-11	Sequence 11, Appl
30	37.2	17.9	1065	2	US-08-461-379A-11	Sequence 11, Appl
31	37.2	17.9	1065	2	US-08-462-390B-11	Sequence 11, Appl
32	37.2	17.9	1065	3	US-08-463-074B-11	Sequence 11, Appl
33	37.2	17.9	1065	3	US-08-465-585C-11	Sequence 11, Appl
34	37.2	17.9	1065	3	US-08-652-446-11	Sequence 11, Appl
35	37.2	17.9	1208	2	US-08-403-852B-4	Sequence 4, Appli
36	37.2	17.9	1208	3	US-08-510-646B-4	Sequence 4, Appli
37	37.2	17.9	1208	4	US-09-231-818-4	Sequence 4, Appli
38	37.2	17.9	5392	2	US-08-403-852B-1	Sequence 1, Appli
39	37.2	17.9	5392	3	US-08-510-646B-1	Sequence 1, Appli
40	37.2	17.9	5392	4	US-09-231-818-1	Sequence 1, Appli
41	37.2	17.9	50937	4	US-09-428-517-1	Sequence 1, Appli
c 42	36.8	17.7	4411529	4	US-09-103-840A-1	Sequence 1, Appli
43	36.6	17.6	829	4	US-09-313-300-7	Sequence 7, Appli
44	36.6	17.6	44377	2	US-08-804-227C-7	Sequence 7, Appli
45	36.6	17.6	44377	2	US-08-804-198-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-250-030-1
: Sequence 1, Application US/08250030
: Patent No. 5643723
: GENERAL INFORMATION:
: APPLICANT: Persing, David H.
: TITLE OF INVENTION: Detection of a Genetic Locus Encoding
: TITLE OF INVENTION: Resistance to Rifampin in Mycobacterial Cultures and in
: TITLE OF INVENTION: Clinical Specimens
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESS: Schwegman, Lundberg & Woessner
: STREET: 3500 IDS Center
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402
COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/250,030
: FILING DATE: 26-MAY-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Mueling, Ann M.
: REGISTRATION NUMBER: 33,977
: REFERENCE/DOCKET NUMBER: 150.105US1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-339-0331
: TELEFAX: 612-339-3061
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 970 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
US-08-250-030-1

Query Match 80.0%; Score 166.4; DB 1; Length 970;
Best Local Similarity 87.5%; Pred. No. 4e-30;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGACTGGCCCGGTAGGCGCTACAAGGTCACAAGAAGTCG 60
|||||
Db 26 TCAAGGAGAGCGCTACGACTGGCCCGGTAGGTCACAAGAAGTCG 85

RESULT 3
US-08-313-185-57
; Sequence 57, Application US/08313185
; Patent No. 5851763
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying

App
add

APPLICANT: Bodmer, Thomas
TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
TITLE OF INVENTION: in Mycobacterium Tuberculosis
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.

```

? STATE: D.C.
? COUNTRY: USA
? ZIP: 20005-3315
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: DOS

```

```
% OPERATING SYSTEM: PC-DOS/MS-DOS
% SOFTWARE: PatentIn Release #1.0, Version #1.25
% CURRENT APPLICATION DATA:
% APPLICATION NUMBER: US/08/313,185
% FILING DATE: 12-OCT-1994
% CLASSIFICATION: 435
% ATTORNEY/AGENT INFORMATION:
% NAME: Meyers, Kenneth J.
% ADDRESS:
```

```

? REGISTRATION NUMBER: 25,146
? REFERENCE/DOCKET NUMBER: 02356.0068-00000
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 408-4000
? TELEFAX: (202) 408-4400
? INFORMATION FOR SEQ ID NO: 57:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 3447 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
?
? US-08-313-185-57
?
? Query Match 68.5%; Score 142.4; DB 2; Length 3447;
? Best Local Similarity 80.3%; Pred. No. 1.4e-24;

```

QY 1 TCAAGGAGAAAGCGCTACGACCTGGCCCGGTAGCCGCTACAGGTTCAACAGAAGCTCG 60

Db	809	TC AAGGAGAAACGGTACGAGCTCGCCAGGGTTGGTCGTTACAAGGTCACAAGAAGCTCG	868
QY	61	GCCTCAGCTCGGGATCCGATCACCAGCTCCACGCTGACCAGGAGACGTCGTCCGCA	120
Db	869	GGTTGCAGCCGGTGAGTTGATCAGCTGCTCCACGCTGACCCAGAGAGATGTGCTGCCA	928
QY	121	CCATCGAGTACTCTGCTCGGCTGCACGAGGGCCACACACGATGACCTGCCGGGGGCA	180
Db	929	CCATAGACTACTGTTCTGCTTCGNTAGGGGTAGTCGACAATGACTGTCCCAGGTGGG	988
QY	181	CCGAGGTGCCGGTTTGAGACCGAGGACAT	208
Db	989	TAGAGATGCCACTCGAAGACATCAATAT	1015

RESULT 4
US-09-082-61AA-57

2-614A-57

; Sequence 57, Application US/09082614A
; Patent No. 6124098
; GENERAL INFORMATION:
; APPLICANT: Heym, Beate
; APPLICANT: Cole, Stewart
; APPLICANT: Young, Douglas
; APPLICANT: Zhang, Ying
; APPLICANT: Honore, Nadine
; APPLICANT: Telenti, Amalio
; APPLICANT: Bodmer, Thomas
; TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
; TITLE OF INVENTION: in Mycobacterium Tuberculosis
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/082.614A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/313,185
; FILING DATE: 12-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 02356.0068-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-082-614A-57

Query Match 68.5%; Score 142.4; DB 3; Length 3447;
Best Local Similarity 80.3%; Pred. No. 1.4e-24;
Matches 167; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGGTACGACCTGGCCCGGTAGGCCGCTACAAGGTCAACAAGAGCTCG 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 809 TCAAGGAGAGACGCTACGACCTGGCCCGGTAGGCCGCTACAAGGTCAACAAGAGCTCG 868
QY 61 GCGTCAGTCGGGATCCGATCCACCTCCACGCTGACCGAGGAGACGCTCGTCGCCA 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 869 GGTTCACCGCGGTGAGTTGATCAGCTGCTCCAGCTGACCGAAGAGAGATGCTCGTCGCCA 928
QY 121 CCATCGAGTACCTGGTCCGCTCGACGAGGCGCCAGCACGATGACCGTCCCGGCGGCA 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 929 CCATAGAGTACCTGGTTCGTGTCATGAGGTGCTGACATGACATGCTCCAGGTGGG 988
QY 181 CCGAGGTCCCGTTGAGACCGACGACAT 208
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 989 TAGAGTGCCAGTGGAACACTGACGATAT 1016

RESULT 5
US-09-029-603-4
; Sequence 4, Application US/09029603

; Patent No. 6210935
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Engel, Natalie
; APPLICANT: Bietenhader, Jurg
; APPLICANT: Toupet, Christine
; APPLICANT: Pospiech, Andreas
; TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters
; FILE REFERENCE: 4-20555/A/PCT
; CURRENT APPLICATION NUMBER: US/09/029,603
; CURRENT FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: PCT/EP96/03643
; EARLIER FILING DATE: 1996-08-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 6085
; TYPE: DNA
; ORGANISM: Streptomyces longisporoflavus
; FEATURE:
; NAME/KEY: misc_RNA
; LOCATION: (378)..(1665)
; OTHER INFORMATION: ORF
; FEATURE:
; NAME/KEY: misc_RNA
; LOCATION: (1747)..(2553)
; OTHER INFORMATION: ORF
; FEATURE:
; NAME/KEY: misc_RNA
; LOCATION: (2593)..(4011)
; OTHER INFORMATION: ORF
; FEATURE:
; NAME/KEY: misc_RNA
; LOCATION: (4013)..(4999)
; OTHER INFORMATION: ORF
; FEATURE:
; NAME/KEY: misc_RNA
; LOCATION: (5071)..(6085)
; OTHER INFORMATION: ORF
US-09-029-603-4

Query Match 23.9%; Score 49.8; DB 4; Length 6085;
Best Local Similarity 60.7%; Pred. No. 0.0022;
Matches 99; Conservative 0; Mismatches 62; Indels 2; Gaps 1;
QY 18 GACCTGGCCCGGTAGGCCGCTACAAGGTCAACAAGAGTCCGCTCGCACGTGCGGAT 77
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1017 GAGGTGGACCGCGCGCGCGGCGGACGACGATGCTCACCCTCTCTCGTCGCGGCC 1076
QY 78 CCATCACCAGTCCACGCTGACCGAGGAGAGAGAGTCCGTCGCCACCATCGAGTACCTGGT- 136
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1077 CGGACACCGGATCAGCGCTCAGCGTGGAGCGCATCGTGGCACCTGCGTCCATCTGCTC 1136
QY 137 -CCGCTTGCACGAGGCGGCGGACGACGATGACCGTCCCGGCGG 178
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1137 ACCGCGGCGGACGAGACCCACCACCACTGCTCGCCAGGCGG 1179
RESULT 6
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match      21.8%; Score 45.4; DB 4; Length 4403765;
Best Local Similarity 52.4%; Pred. No. 0.036;
Matches 100; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy 10 AGCGCTAGCACCTGGCCGGGTAGGCGCTACAAAGGTCAACAAGAAGCTCGGCGCTGCAGC 69
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4013310 AGCAACCGCATAGTTCAGGTGGTTCGCGATGTTGACCCCGAGTAGCAGCCGACGAGC 4013251

Qy 70 TCGCGGATCCGATCACCAGCTCCACGCTGACCGAGGAAGACGTGCTGCCACCATCGAGT 129
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4013250 GCCGCGCGCGCGCGCGCCGCCACCGGCGCACTAACACAGAGTCGCGGGCAGGTGTG 4013191

Qy 130 ACTGCTCCGCTCCACAGGGCCAGCACAGCATGACCGTCCCGGGCGGCGACCGAGGTGC 189
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4013190 ACCACGTTGGCCAGCACGCGGGCCAGCGCGCATGCCGAGCAGCGGGGCAACCCCGGACC 4013131

Qy 190 CGGTTGAGACC 200
      ||||| |||||
Db 4013130 CGGTCGGCAGC 4013120

RESULT 7
US-09-029-603-1
; Sequence 1, Application US/09029603
; Patent No. 6210935
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Engel, Natalie
; APPLICANT: Bietenhader, Jurg
; APPLICANT: Toupet, Christine
; APPLICANT: Pospiech, Andreas
; TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters
; FILE REFERENCE: 4-20555/A/PCT
; CURRENT APPLICATION NUMBER: US/09/029,603
; CURRENT FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: PCT/EP96/03643
; EARLIER FILING DATE: 1996-08-19
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2122
; TYPE: DNA
; ORGANISM: Streptomyces longisporoflavus
; FEATURE:
; NAME/KEY: misc_RNA
; LOCATION: (1)..(2122)
; OTHER INFORMATION: product = 2.1 kb region
US-09-029-603-1

Query Match      20.7%; Score 43; DB 4; Length 2122;
Best Local Similarity 52.5%; Pred. No. 0.073;
Matches 94; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy 8 GAAGCGGTACAGCTGGCCCGGGTAGCGCGCTACAAGGTCAACAAGAAGCTCGGCGTGA 67
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1470 GCACCGCTCCGGGTGCGGGTGCAGAGATCGTCAGCTACCGGAGCAGACAGCTGCG 1529

Qy 68 CGTGGGCGATCCGATCACCAGCTCCAGCTGACCGAGGAAGAGTCTCGCCACCATCGA 127
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1530 CTTCCGCGACGAGATCGGCAAGCTCGCGCGCGTTCGAGGAGCGCCCCGTGGCCATGGACGA 1589

Qy 128 GTACCTGCTCGCTGCACGAGGCGCAGCACACGATGACCGTCCCGGGCGGCGACCGAGG 186
```

```

Db 1590 GGGCAACTTCGCTTCGGCGGACGACTCCTTCAAGCGCTCCGACCTGGCGGCGTGCAGC 1648

RESULT 8
US-08-476-519-10
; Sequence 10, Application US/08476519
; Patent No. 5750876
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Krohn, Bradley M.
; TITLE OF INVENTION: No. 5750876el Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
; STREET: 700 Chesterfield Parkway No. 5750876th
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,519
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/281902
; FILING DATE: 28-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(13577)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-7286
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2244 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2241
US-08-476-519-10

Query Match      20.2%; Score 42; DB 1; Length 2244;
Best Local Similarity 52.9%; Pred. No. 0.12;
Matches 90; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

Qy 17 CGACCTGGCCCGGGTAGGCGCGTACAAGGTCAACAAGAAGCTCGGCGCTGCAGCTCGGCGA 76
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1338 CGAGTGAACGCGCTCTACCGCGACGCGCTGCGCAAGAGAGCAACAAGCTCGGCGTGA 1397

Qy 77 TCGGATCACAGCTCCACGCTGACCGAGGAGACAGCTGCTGCGCCACCATCGAGTACCTGGT 136
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1398 AAGGTCACCGCGGACCGCTTCCGCGGCTCCACGACCTGCTGCGCCGCGA 1457

Qy 137 CGCGCTGCACGAGGCGCAGCACGATGACCTCCCGGGCGGCGACCGAGG 186
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1458 CGAGGCGCCCAAGCCGTCATTCGATCACTTCGTTGTCGCCCGAGGAG 1507

RESULT 9
PCT-US95-09323-10
; Sequence 10, Application PC/TUS9509323
```

```

: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Novel Isoamylase Gene, Compositions
: CONTAINING IT AND METHODS OF USING ISOAMYLASES
: NUMBER OF SEQUENCES: 11
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: FILING DATE:
: APPLICATION NUMBER: PCT/US95/09323
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/281902
: FILING DATE: 28-JUL-1994
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2244 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..2241
: PCT-US95-09323-10

Query Match 20.2%; Score 42; DB 5; Length 2244;
Best Local Similarity 52.9%; Pred. No. 0.12;
Matches 90; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 17 CGACCTGCCCCGGGTAGCCCTACAGGTCAACAAGAGCTCGGCTGACGTGCGCGA 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1338 CGAGTGAACGGCTCTACCGCGCGCTGCGCAAGAGCAGCAAGCTCGCGGTGA 1397

QY 77 TCCGATCACAGCTCCAGCTGACCGAGGAGACGTCGTCGCCACCATCGACTGGT 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1398 AACGCTACCCCGGACCCCTGGCCAGCGCTTCGCGGCTCCCAACGACCTGTACGCGGA 1457

QY 137 CGCGCTGCACGAGGCCAGCACACGATGACGTCCTCCGGGGCGGACCGAGG 186
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1458 CGAGCGCGCAAGCGGTGGTCACTTCGATCACTTCGTCGTCGCCACGACG 1507

RESULT 10
US-08-476-519-1
: Sequence 1, Application US/08476519
: Patent No. 5750876
: GENERAL INFORMATION:
: APPLICANT: Barry, Gerard F.
: APPLICANT: Kishore, Ganesh M.
: TITLE OF INVENTION: Novel Isoamylase Gene, Compositions
: CONTAINING IT AND METHODS OF USING ISOAMYLASES
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
: STREET: 700 Chesterfield Parkway No. 5750876th
: CITY: St. Louis
: STATE: Missouri
: COUNTRY: USA
: ZIP: 63198
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: FILING DATE:
: APPLICATION NUMBER: US/08/476,519
: CLASSIFICATION: 800

```

```

: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/281902
: FILING DATE: 28-JUL-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Bonner, Grace L.
: REGISTRATION NUMBER: 32,963
: REFERENCE/DOCKET NUMBER: 38-21(13577)A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (314)537-7286
: TELEFAX: (314)537-6047
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2334 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-476-519-1

Query Match 20.2%; Score 42; DB 1; Length 2334;
Best Local Similarity 52.9%; Pred. No. 0.13;
Matches 90; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 17 CGACCTGGCCCGGTAGCGCTACAAAGGTCAACAAGAGCTCGGCTGACGTGCGCGA 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1431 CGAGTGAACGGCTCTACCGCGCGCTGCGCAAGAGCAGCAAGCTCGGCGTGA 1490

QY 77 TCCGATCACAGCTCCAGCTGACCGAGGAGACGTCGTCGCCACCATCGACTGGT 136
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1491 AACGGTCACCCCGGACCCCTGGCCACGCGCTTCGCGGCTCCCAACGACCTGTACGCGGA 1550

QY 137 CGCGCTGCACGAGGCCAGCACACGATGACGTCCTCCGGGGCGGACCGAGG 186
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1551 CGACGGCGCAAGCGGTGGATTCGATCACTTCGTCGTCGCCACGAGG 1600

RESULT 11
PCT-US95-09323-1
: Sequence 1, Application PC/TUS9509323
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: Novel Isoamylase Gene, Compositions
: CONTAINING IT AND METHODS OF USING ISOAMYLASES
: NUMBER OF SEQUENCES: 11
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/09323
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/281902
: FILING DATE: 28-JUL-1994
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2334 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: PCT-US95-09323-1

Query Match 20.2%; Score 42; DB 5; Length 2334;
Best Local Similarity 52.9%; Pred. No. 0.13;
Matches 90; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 17 CGACCTGGCCCGGTAGCGCTACAAAGGTCAACAAGAGCTCGGCTGACGTGCGCGA 76
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1431 CGAGTGAACGGCTCTACCGCGCGCTGCGCAAGAGCAGCAAGCTCGGCGTGA 1490

```

```
QY 77 TCCGATCACCAGCTCCACGCTGACCGAGGAAGACGCTCGTCGCCACCATCTGAGTACCTGGT 136
      ||||| | ||||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 1491 AACGTCACCCCGCCACCCCTGGCCACCGCGCTTCCCGCGCTCCACAGACCTGTACGGCGA 1550
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 137 CCGCTGTCAGGAGGCCAGCAGCAGATGACCGTCCCGGGCGGCACCCGAGG 186
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1551 CGACGGCCGCAAGCGGTGGCATTCGATCAACTTCGTGTGTCGCCACGACG 1600
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-09-103-840A-2
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 19.3%; Score 40.2; DB 4; Length 4403765;
Best Local Similarity 51.4%; Pred. No. 0.55;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 21 CTGCGCGGGTAGGCGGTACAAAGTCAACAAGAGTCGGCTGCGGATCGGCGGATCCG 80
      ||||| ||| | | | | | | | | | | | | | | | | | | | | | | |
Db 175075 CTGCGCGGGAGGGCGCCAGCGCTGTTGTTCAACGACCTCGTGGCGCCCGCGGCGACG 175134
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 81 ATCACCAGCTCCACGCTGACCGAGGAAGAGGTGTCGCCACCATCTGCTGCTCCG 140
      ||||| ||| | | | | | | | | | | | | | | | | | | | | | |
Db 175135 GCGCGCGGTTGCGCGATGCGCGATGAGTGTGTCGCCGAGATTCGCGCAAGGGGCGG 175194
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 141 CTGACAGGAGGCCAGCAGCAGATGACCGTCCCGGGCGGCACCGAGGTGCCGTTGAGACC 200
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 175195 GCGGTGCGCAACTACGACAGCGTCGCCACCGAGGACGCGCAGCAACATCATCAAGACC 175254
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 201 G 201
Db 175255 G 175255

RESULT 13
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
```

```
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 19.3%; Score 40.2; DB 4; Length 4411529;
Best Local Similarity 51.4%; Pred. No. 0.55;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 21 CTGCGCGGGTAGGCGGTACAAAGTCAACAAGAGTCGGCTGCGGATCGGCGGATCCG 80
      ||||| ||| | | | | | | | | | | | | | | | | | | | | | |
Db 174906 CTGCGCGGGAGGGCGCCACCGCTCGTGGTCAACGACCTCGGTGCGCCCGCGGACGCGCAG 174965
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 81 ATCACCAGTCCACGCTGACCGAGGAAGAGGTGTCGCCACCATCTGAGTACCTGTTGTCGCG 140
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 174966 GCGCGCGGTTGCGGATGCGCGATGAGTGTGTCGCCGAGATTCGCGCAAGGGGCGCGG 175025
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 141 CTGACAGGAGGCCAGCAGCAGATGACCGTCCCGGGCGGCACCGAGGTGCCGTTGAGACC 200
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 175026 GCGTTCGCCAACTACGACAGCGTCCCGACCGAGGACGCGCGCAACATCATCAAGACC 175085
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 201 G 201
Db 175086 G 175086

RESULT 14
US-08-387-942C-1
; Sequence 1, Application US/08387942C
; Patent No. 5939289
; GENERAL INFORMATION:
; APPLICANT: ERTESVAG, HELGA
; APPLICANT: VALLA, SVEIN
; APPLICANT: SKJAK-BRAEK, GUDMUND
; APPLICANT: LARSEN, BJORN
; TITLE OF INVENTION: DNA COMPOUNDS COMPRISING SEQUENCES
; TITLE OF INVENTION: ENCODING MANNURONAN C-5-EPIMERASE
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O. BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/387,942C
; FILING DATE: 09-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1809-106P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12588 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Azotobacter vinelandii
; STRAIN: E
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: 290..1951
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2227..6438
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 6702..9695
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 9973..12588
; US-08-387-942C-1

Query Match      18.8%; Score 39.2; DB 2; Length 12588;
Best Local Similarity 50.58; Pred. No. 0.62; Indels 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 17 CGACCTGCGCCGGTAGCCGCTACAAAGTCAACAAGAAGCTCGCCCTGCACGTCGGCGA 76
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3771 CGCCCTGGACGGCAACTTCACCGGTCTGTTCAACGACAACAACCTGTTGTCGACGCCG 3830
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 77 TCCGATCACCAGCTCCAGCTGACGGAGGAAGACGTGTCGCCACCATCGAGTACCTGGT 136
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3831 TCCGCCACACCGGTACCGAGGGCAGCAACCTGCTCGGCACCGCGGGGAAACCT 3890
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 137 CGCCTGCAGGAGGCCAGCAGATGACCGTCCCGGGCGGCACCGAGGTGCCGGTTGA 196
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3891 CTGGGCTACGGCGCAACGACACCTCAACGGCGGGCGCGGACGACATCCTGGTCGG 3950
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 197 GACCGAGC 204
   || ||
Db 3951 CGGCGCG 3958

RESULT 15
US-08-804-227C-7/c
; Sequence 7, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rosteck, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4437 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 350..14002
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31329..36071
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36155..41830
; US-08-804-227C-7

Query Match      18.8%; Score 39; DB 2; Length 44377;
Best Local Similarity 53.6%; Pred. No. 0.75; Indels 0; Gaps 0;
Matches 81; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 2 CAAGGAGAGCGCTACGACCTGGCCCGGTAGGCCGCTACAAAGTCAACAAGAAGCTCGG 61
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 43485 CGAGGAGCAGCGCGATGCGCTCGACGTCGCCGGGGCGGACGCCGTCGTGCGAGGAGGG 43426
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 62 CCTGCACGTCGGCGATCCGATCACCAGCTCCACGCTGACCGAGGAAAGACGTCGTGCCAC 121
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 43425 CGAGCAGCTGTCGGCGCGGGGCTGCCGCTTCCGAGGCGCCGGGACTTCTCCGCGACGA 43366
   ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 122 CATCGAGTACCTGGTCGGCTGCACAGGGC 152
   || | || | || | || | || | || | || | || | || | || | || | || | || | || |
Db 43365 GCTCGGGGACCAGGTGCGCCAGTGTGCGGAC 43335
   || | || | || | || | || | || | || | || | || | || | || | || | || | || |

Search completed: November 12, 2002, 18:08:49
Job time : 4510.89 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 ; Search time 136.222 Seconds
(without alignments)
3438.621 Million cell updates/sec

Title: us-09-697-123b-1

Perfect score: 208
Sequence: 1 tcaagagagacgctacgac.....ccggtgagaccgacacat 208

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	208	100.0	208	22	AAS05201 Mycobacterium gord
2	195.2	93.8	208	22	AAS05203 Mycobacterium gord
3	190.4	91.5	208	22	AAS05215 Mycobacterium avi
4	177.6	85.4	208	22	AAS05202 Mycobacterium gord
5	177.6	85.4	208	22	AAS05211 Mycobacterium mari
6	175.2	84.2	207	22	AAS05204 Mycobacterium gord
7	172.8	83.1	208	22	AAS05210 Mycobacterium ulce
8	169.6	81.5	208	22	AAS05219 Mycobacterium fort
9	168	80.8	205	22	AAS05220 Mycobacterium intr

10	168	80.8	208	22	AAS05208 Mycobacterium kans
11	166.4	80.0	208	22	AAS05205 Mycobacterium tube
12	166.4	80.0	208	22	AAS05206 Mycobacterium terr
13	166.4	80.0	208	22	AAS05216 Mycobacterium bovi
14	166.4	80.0	208	22	AAS05217 Mycobacterium ceta
15	166.4	80.0	970	17	AAT09676 Mycobacterium tube
16	166.4	80.0	3519	22	AAH51976 Mycobacterium tube
17	166.4	80.0	3534	22	AAH02079 Mycobacterium tube
18	166.4	80.0	3853	21	AAA74651 Mycobacterium tube
19	166.4	80.0	3853	21	AAA89994 M. tuberculosis rp
20	164.8	79.2	208	22	AAS05222 Mycobacterium afri
21	164.8	79.2	208	22	AAS05223 Mycobacterium haem
22	164	78.8	207	22	AAS05212 Mycobacterium szul
23	155.2	74.6	208	22	AAS05224 Mycobacterium xeno
24	142.4	68.5	214	22	AAS05221 Mycobacterium abs
25	142.4	68.5	3447	14	AAQ51532 M. leprae rpoB gene
26	140.8	67.7	214	22	AAS05214 Mycobacterium malm
27	139.8	67.2	223	22	AAS05209 Mycobacterium scro
28	139.2	66.9	214	22	AAS05207 Mycobacterium chel
29	135	64.9	223	22	AAS05213 Mycobacterium gast
30	131	63.0	211	22	AAS05218 Mycobacterium flav
31	83.2	40.0	3495	22	AAH65512 C glutamicum codin
32	83.2	40.0	34980	22	AAH68525 C glutamicum codin
33	73.6	35.4	27426	23	AAH59541 Propionibacterium
34	49.8	23.9	6085	18	AAT70153 S. longisporoflavus
35	45.4	21.8	4403765	22	AA199863 Mycobacterium tube
36	45.2	21.7	2104	21	AAAC75978 Human ORFX ORF1533
37	44	21.2	21185	21	AAA63350 Streptomyces globi
38	44	21.2	63164	21	AAA63348 Streptomyces globi
39	43	20.7	2122	18	AAT70152 S. longisporoflavus
40	42.2	20.3	77536	21	AAA14651 Nucleotide sequenc
41	42	20.2	1844	12	AAQ12000 Maize nitrite redu
42	42	20.2	2244	17	AAT10429 Mature isoenzyme
43	42	20.2	2244	19	AAV23640 Flavobacterium iso
44	42	20.2	2334	17	AAT10428 Isoamylase gene.
45	42	20.2	2334	19	AAV23639 Full length Flavob

ALIGNMENTS

RESULT 1

AAS05201

ID AAS05201 standard; DNA; 208 BP.

AC AAS05201;

XX

XX 07-SEP-2001 (first entry)

DT

DE Mycobacterium gordonae type I rpoB gene fragment.

XX

KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

KW PCR-restriction fragment length polymorphism analysis; ds.

XX

OS Mycobacterium gordonae type I.

XX

BN WO200131061-A1.

XX

PD 03-MAY-2001.

XX

PF 27-OCT-2000; 2000WO-KR01223.

XX

PR 27-OCT-1999; 99KR-0046795.

XX

PA (ERUM-) ERUME BIOTECH CO LTD.

XX

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

```
XX Claim 1; Page 40; 50pp; English.
XX
XX The present sequence for Mycobacterium gordonae type I rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 46 A; 71 C; 65 G; 26 T; 0 other;
XX
XX Query Match 100.0%; Score 208; DB 22; Length 208;
XX Best Local Similarity 100.0%; Pred. No. 8.9e-36;
XX Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TCAAGGAGAACGCGCTACGACCTGGCCGCGGTAGGCGCTACAAGGTCAACAAGAAGCTCG 60
DB 1 TCAGGAGAACGCGCTACGACCTGGCCGCGGTAGGCGCTACAAGGTCAACAAGAAGCTCG 60
QY 61 GCCTGACGCTGGCGATCGGATCACCAGCTCCAGCTGACCGAGGAGACGCTGTCGCCA 120
DB 61 GCCTGACGCTGGCGATCGGATCACCAGCTCCAGCTGACCGAGGAGACGCTGTCGCCA 120
QY 121 CCATCGAGTACCTGGTCCGCTGCAGGGCCAGCACACGATGACCGTCCCGGGCGGCA 180
DB 121 CCATCGAGTACCTGGTCCGCTGCAGGGCCAGCACACGATGACCGTCCCGGGCGGCA 180
QY 181 CCGAGGTGCGCGTTGAGACCGACGACAT 208
DB 181 CCGAGGTGCGCGTTGAGACCGACGACAT 208
XX
RESULT 2
AAS05203
ID AAS05203 standard; DNA; 208 BP.
XX
XX AAS05203;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium gordonae type III rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
XX Mycobacterium gordonae type III.
XX
XX WO200131061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
```

```
XX WPI; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
XX Claim 1; Page 41; 50pp; English.
XX
XX The present sequence for Mycobacterium gordonae type III rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 46 A; 69 C; 63 G; 30 T; 0 other;
XX
XX Query Match 93.8%; Score 195.2; DB 22; Length 208;
XX Best Local Similarity 96.2%; Pred. No. 4.7e-33;
XX Matches 200; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
XX
QY 1 TCAAGGAGAACGCGCTACGACCTGGCCGCGGTAGGCGCTACAAGGTCAACAAGAAGCTCG 60
DB 1 TCAAGGAGAACGCGCTACGACCTGGCCGCGGTAGGCGCTACAAGGTCAACAAGAAGCTCG 60
QY 61 GCCTGACGCTGGCGATCGGATCACCAGCTCCAGCTGACCGAGGAGACGCTGTCGCCA 120
DB 61 GCCTGACGCTGGCGATCGGATCACCAGCTCCAGCTGACCGAGGAGACGCTGTCGCCA 120
QY 121 CCATCGAGTACCTGGTCCGCTGCAGGGCCAGCACACGATGACCGTCCCGGGCGGCA 180
DB 121 CCATCGAGTACCTGGTCCGCTGCAGGGCCAGCACACGATGACCGTCCCGGGCGGCA 180
QY 181 CCGAGGTGCGCGTTGAGACCGACGACAT 208
DB 181 CCGAGGTGCGCGTTGAGACCGACGACAT 208
XX
RESULT 3
AAS05215
ID AAS05215 standard; DNA; 208 BP.
XX
XX AAS05215;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium avium rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
XX Mycobacterium avium.
XX
XX WO200131061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX
```

```
XX 27-OCT-1999; 99KR-0046795.
XX (ERUM-) ERUME BIOTECH CO LTD.
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of Mycobacterium, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism -
XX
XX Claim 1; Page 44; 50pp; English.
XX
XX The present sequence for Mycobacterium avium rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;
XX
XX Query Match 91.5%; Score 190.4; DB 22; Length 208;
XX Best Local Similarity 94.7%; Pred. No. 5e-32;
XX Matches 19; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
XX
XX 1 TCAAGGAGAGCGCTACGACCTGCGCCGGGTAGCGGCTACAAAGGTCAACAAGAGCTCG 60
XX
XX 1 TCAAGGAGAGCGCTACGACCTGCGCCGGGTAGCGGCTACAAAGGTCAACAAGAGCTCG 60
XX
XX 61 GCCTGCGAGTCCGCGATCCGATCACAGCTCCAGCTGACCGAGGAGAGCTCGTCCGCA 120
XX
XX 61 GCCTGCGAGTCCGCGATCCGATCACAGCTCCAGCTGACCGAGGAGAGCTCGTCCGCA 120
XX
XX 121 CCATCGAGTACCTGCTCGGCTGCAGGAGGCGCCAGCAGCATGACCGTCCCGGGCGGCA 180
XX
XX 121 CCATCGAGTACCTGCTCGGCTGCAGGAGGCGCCAGCAGCATGACCGTCCCGGGCGGCA 180
XX
XX 181 CCGAGGTGCGGGTGTGAGACCGACGACAT 208
XX
XX 181 TCGAGGTGCGGGTGTGAGACCGACGACAT 208
XX
XX
XX RESULT 4
XX AAS05202
XX ID AAS05202 standard; DNA; 208 BP.
XX
XX AC AAS05202;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium gordonae type II rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
XX Mycobacterium gordonae type II.
```

```
XX WO200131061-A1.
XX 03-MAY-2001.
XX 27-OCT-2000; 2000WO-KR01223.
XX 27-OCT-1999; 99KR-0046795.
XX (ERUM-) ERUME BIOTECH CO LTD.
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism -
XX
XX Claim 1; Page 40; 50pp; English.
XX
XX The present sequence for Mycobacterium gordonae type II rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 45 A; 62 C; 73 G; 28 T; 0 other;
XX
XX Query Match 85.4%; Score 177.6; DB 22; Length 208;
XX Best Local Similarity 90.9%; Pred. No. 2.7e-29;
XX Matches 189; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
XX
XX 1 TCAAGGAGAGCGCTACGACCTGCGCCGGGTAGCGGCTACAAAGGTCAACAAGAGCTCG 60
XX
XX 1 TCAAGGAGAGCGCTACGACCTGCGCCGGGTAGCGGCTACAAAGGTCAACAAGAGCTCG 60
XX
XX 61 GCCTGCGAGTCCGCGATCCGATCACAGCTCCAGCTGACCGAGGAGAGCTCGTCCGCA 120
XX
XX 61 GTCGAACTGCGGAGCGGATCCAGCTCCAGCTGACCGAGGAGAGCTCGTCCGCA 120
XX
XX 121 CCATCGAGTACCTGCTCGGCTGCAGGAGGCGCCAGCAGCATGACCGTCCCGGGCGGCA 180
XX
XX 121 CCATCGAGTACCTGCTCGGCTGCAGGAGGCGCCAGCAGCATGACCGTCCCGGGCGGCA 180
XX
XX 181 CCGAGGTGCGGGTGTGAGACCGGACGACAT 208
XX
XX 181 CCGAGGTGCGGGTGTGAGACCGGACGACAT 208
XX
XX
XX RESULT 5
XX AAS05211
XX ID AAS05211 standard; DNA; 208 BP.
XX
XX AC AAS05211;
XX
XX 07-SEP-2001 (first entry)
XX
```



```

Db 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGCGCGCTACAAAGAGCTGG 60
Qy 61 GCCTGCACGTCGGCGATCACCAGCTCACGCTGACCGAGGAAGACGTCGTCGCCA 120
Db 61 GCCTGAACCGCGCCAGCGGATCAGCTCGTCTGACCGAGGAAGACGTCGTCGCCA 120
Qy 121 CCATCCAGTACCTGGTCCGCTGCAGGAGGCCAGCACACGATGACCGTCCCGGCGGCCA 180
Db 121 CCATCCAGTACCTGGTCCGCTGCAGGAGGCCAGCACACGATGACCGTCCCGGCGGCCA 180
Qy 181 CCGAGTGCCGGTTGAGACCGACGACAT 208
Db 181 TCGAGTCCCGTTCGAGGTGGACGACAT 208

RESULT 9
AAS05220
ID AAS05220 standard; DNA; 205 BP.
XX
AC AAS05220;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium intracellulare rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium intracellulare.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
XX WPI; 2001-300520/31.
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
PS Claim 1; Page 46; 50pp; English.
XX
CC The present sequence for Mycobacterium intracellulare rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 205 BP; 44 A; 69 C; 67 G; 25 T; 0 other;

```

```

Query Match 80.8%; Score 168; DB 22; Length 205;
Best Local Similarity 91.3%; Pred. No. 3e-27;
Matches 190; Conservative 0; Mismatches 15; Indels 3; Gaps 1;

Qy 1 TCAAGGAGAGCGCTACGACCTGGCCCGGGTAGCCGCTACAAAGTCAACAAGAAGCTCG 60
Db 1 TCAAGGAGAGCGCTACGACCTGGCCCGGGTAGCCGCTACAAAGTCAACAAGAAGCTCG 60
Qy 61 GCCTGCACGTCGGCGATCACCAGCTCACGCTGACCGAGGAAGACGTCGTCGCCA 120
Db 61 GCCTGCACGTCGGCGATCACCAGCTCACGCTGACCGAGGAAGACGTCGTCGCCA 120
Qy 121 CCATCCAGTACCTGGTCCGCTGCAGGAGGCCAGCACACGATGACCGTCCCGGCGGCCA 180
Db 121 CCATCCAGTACCTGGTCCGCTGCAGGAGGCCAGCACACGATGACCGTCCCGGCGGCCA 180
Qy 181 CCGAGTGCCGGTTGAGACCGACGACAT 208
Db 178 TCGAGGTCCCGTTCGAGGTGGACGACAT 205

```

```

RESULT 10
AAS05208
ID AAS05208 standard; DNA; 208 BP.
XX
AC AAS05208;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium kansasii rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium kansasii.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
XX WPI; 2001-300520/31.
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
PS Claim 1; Page 42; 50pp; English.
XX
CC The present sequence for Mycobacterium kansasii rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR

```

CC fragments with the unidentified fragment. The rpoB gene fragments


```

DR WPI; 2001-300520/31.
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism
XX
PS Claim 1; Page 45; 50pp; English.
XX
CC The present sequence for Mycobacterium celatum rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 48 A; 71 C; 64 G; 25 T; 0 other;

Query Match 80.0%; Score 166.4; DB 22; Length 208;
Best Local Similarity 87.5%; Pred. No. 6.5e-27;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGCTACGACCTGGCCGGGTAGGCGCTACAAAGGTCAACAAGAGCTCG 60
Db 1 TCAAGGAGAAGCGCTACGACCTGGCCGGGTAGGCGCTACAAAGGTCAACAAGAGCTCG 60
QY 61 GCTCTGACGTCGCGGATCCGATCACCAGCTCCACGCTGACCGAGGAGACGTCGTGCCA 120
Db 61 GCTCTGACACCGGCTCCCGATCCAGCAGCCACTCTGACCGAAGAGGACGTCGTGCCA 120
QY 121 CCATCGAGTACCTGGTCCGCTGCACGAGGGCCAGCACACGATGACCGTCCCGGGCGGCA 180
Db 121 CCATCGAGTACCTGGTCCGCTGCACGAGGGCCAGCACACGATGACCGTCCCGGGCGGAG 180
QY 181 CCGAGTGCCGGTTGAGACCGAGCAT 208
Db 181 TCGAGGTGCGGTGGAAACCGAGCAT 208

RESULT 15
AAT09676
ID AAT09676 standard; DNA; 970 BP.
AC AAT09676;
XX
DT 15-OCT-1996 (first entry)
XX
DE Mycobacterium tuberculosis rpoB gene DNA sequence.
XX
KW Tuberculosis; disease diagnosis; oligonucleotide; DNA primer; PCR;
KW polymerase chain reaction; DNA amplification; rpoB locus; TB; ss.
XX
OS Mycobacterium tuberculosis.
XX
FH Key Location/Qualifiers
FT primer_bind 10..27
FT /*tag= a
FT /*note= "primer FENLFF"
FT primer_bind 226..243
FT /*tag= b

```

```

FT primer_bind /note= "primer DDIDHL"
FT 226..240
FT /*tag= c
FT primer_bind /note= "primer DDIDH"
FT 338..364
FT /*tag= d
FT primer_bind /note= "primer rpo95"
FT 348..373
FT /*tag= e
FT primer_bind /note= "primer rpo105"
FT 354..373
FT /*tag= f
FT misc_feature /note= "primer KY290"
FT 372..373
FT /*tag= g
FT misc_feature /note= "M. tuberculosis signature nucleotide"
FT 433..434
FT /*tag= h
FT misc_feature /note= "M. tuberculosis signature nucleotide"
FT 438
FT /*tag= i
FT misc_feature /note= "M. tuberculosis signature nucleotide"
FT 468..469
FT /*tag= j
FT misc_feature /note= "M. tuberculosis signature nucleotide"
FT 486
FT /*tag= k
FT misc_feature /note= "M. tuberculosis signature nucleotide"
FT 501
FT /*tag= l
FT misc_feature /note= "M. tuberculosis signature nucleotide"
FT 516
FT /*tag= m
FT primer_bind /note= "M. tuberculosis signature nucleotide"
FT 516..535
FT /*tag= n
FT misc_feature /note= "primer rpo273"
FT 525
FT /*tag= o
FT primer_bind /note= "M. tuberculosis signature nucleotide"
FT 525..541
FT /*tag= p
FT primer_bind /note= "primer KY292"
FT 536..562
FT /*tag= q
FT primer_bind /note= "primer rpo293"
FT 640..666
FT /*tag= r
FT primer_bind /note= "primer rpo397"
FT 952..966
FT /*tag= s
FT primer_bind /note= "primer NMQRQ-1"
FT 952..966
FT /*tag= t
FT primer_bind /note= "primer NMQRQ-2"
FT 952..966
XX WO9533074-A1.
XX
XX 07-DEC-1995.
XX
XX 26-MAY-1995; 95WO-US06790.
XX
XX 26-MAY-1994; 94US-0250030.
XX (HOFF ) HOFFMANN LA ROCHE INC.
XX (MAYO-) MAYO FOUNDATION.
XX
XX Felmlee TA, Hunt JM, Persing DH, Roberts GD, Whelen AC;
XX Young KK;
XX
XX WPI; 1996-030581/03.
XX
XX Detection of Mycobacterium tuberculosis - by amplifying sample DNA
PT

```


GenCore Version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:50:46 ; Search time 636.023 Seconds
(without alignments)
9517.553 Million cell updates/sec

Title: US-09-697-123B-1
Perfect score: 208
Sequence: 1 tcaaggagagcgctacgac.....ccggttgagaccgacgacat 208

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_un:*
- 28: em_vl:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sv:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	166.4	80.0	610	1	MTU318818	AJ318818 Mycobacte
2	166.4	80.0	610	1	MTU318819	AJ318819 Mycobacte
3	166.4	80.0	616	1	MTU318813	AJ318813 Mycobacte
4	166.4	80.0	618	1	MTU318815	AJ318815 Mycobacte
5	166.4	80.0	618	1	MTU318817	AJ318817 Mycobacte
6	166.4	80.0	633	1	MTU318814	AJ318814 Mycobacte
7	166.4	80.0	637	1	MTU318816	AJ318816 Mycobacte
8	166.4	80.0	639	1	MTU318821	AJ318821 Mycobacte
9	166.4	80.0	970	6	IS0706	Sequence 1
10	166.4	80.0	3534	6	AX111339	Sequence
11	166.4	80.0	3853	1	MTU12205	U12205 Mycobacteri
12	166.4	80.0	5084	1	MSGRPOB	L27989 Mycobacteri
13	166.4	80.0	19352	1	AE006964	AE006964 Mycobacte
14	166.4	80.0	19770	1	MTCI376	295972 Mycobacteri
15	160	76.9	3752	1	MSU24494	U24494 Mycobacteri
16	142.4	68.5	3447	6	AR067447	AR067447 Sequence
17	142.4	68.5	37617	1	MLB1790G	214314 M.leprae ge
18	142.4	68.5	37617	1	MLB1790G	AL583923 Mycobacte
19	122	58.7	3941	1	MLEPRTN7	AF242549 Amycolato
20	91.6	44.0	32923	1	SCD82	AL160431 Streptomy
21	83.2	40.0	3495	6	AX120631	AX120631 Sequence
22	83.2	40.0	328050	1	AP005275	AP005275 Corynebac
23	83.2	40.0	349980	6	AX127144	AX127144 Sequence 4
24	49.8	23.9	6085	6	A60304	Sequence 4
25	49.8	23.9	6085	6	ARI44763	ARI44763 Sequence
26	46.2	22.2	12112	1	AE005722	AE005722 Caulobact
27	46.2	22.2	347660	1	AP002994	AP002994 Mesorhizo
28	45.8	22.0	10846	1	AE005837	AE005837 Caulobact
29	45.8	22.0	329100	1	SME591787	AL591787 Sinorhizo
30	45.2	21.7	2100	9	AK096697	AK096697 Homo sapi
31	45.2	21.7	118312	9	AC055713	AC055713 Homo sapi
32	45	21.6	3137	8	AF268091	AF268091 Chloris g
33	44.2	21.3	7854	1	TAQ19223	Y19223 Thermus aqu
34	44.2	21.3	12523	1	AE001898	AE001898 Deinococc
35	44.2	21.3	18160	1	SC6C5	AL034492 Streptomy
36	44.2	21.3	20054	1	SCCB12	AL591588 Streptomy
37	44.2	21.3	134816	2	CNS08C9R	AL772413 Oryza sat
38	44.2	21.3	135378	2	CNS07XP9	AL7713941 Oryza sat
39	44.2	21.3	196050	1	AL646058	AL646058 Ralstonia
40	44	21.2	299350	1	SME591786	AL591786 Sinorhizo
41	43.8	21.1	2966	8	AF247649	AF247649 Conlathyr
42	43.8	21.1	13818	1	AE004747	AE004747 Pseudomon
43	43.8	21.1	33095	1	SC3C8	AL023861 Streptomy
44	43.6	21.0	127108	8	OSJN00144	AL562949 Oryza sat
45	43.6	21.0	130014	2	OSJN00214	AL663013 Oryza sat

ALIGNMENTS

RESULT 1
MTU318818

LOCUS

DEFINITION
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1415-97.

ACCESSION
AJ318818

VERSION
AJ318818.1 GI:22208412

KEYWORDS
RNA polymerase beta subunit; rpoB gene.

SOURCE
Mycobacterium tuberculosis.

ORGANISM
Mycobacterium tuberculosis

REFERENCE
1
Herrera, L., Jimenez, M.S. and Saez, J.A.

```
TITLE
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL
REFERENCE
2 (bases 1 to 610)
AUTHORS
Herrera, L.
TITLE
Direct Submission
JOURNAL
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source
1..610
/organism="Mycobacterium tuberculosis"
/isolate="1415-97"
/db_xref="taxon:1773"
gene
1..610
/genes="rpoB"
<1..>610
/genes="rpoB"
CDS
1..610
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87035.1"
/db_xref="GI:22208413"
/translation="LDIYRKLPGEPPTKESAOITLLENLFFKEKRYDLARVGRYKVNK
KLGHLVGEPIITSLTEDVDVATIEYLRHLEGQTTMTVPGGVEVPVETDDIDHFGNR
RLRTVGELIQNOIRVGMSRMERVMRTQDVEAITPQTLINIRPVAAIKEFFGTS
QPSQFMQNNPLSLGTHKRRLSALGPGGLSRERAGLEVRDHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match 80.0%; Score 166.4; DB 1; Length 610;
Best Local Similarity 87.5%; Pred. No. 1.5e-16;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGCGCGCTACAGGTCACAGAAGCTCG 60
Db 80 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGCGCGCTACAGGTCACAGAAGCTCG 139
QY 61 GCCTGCACCTGGCGATCCAGCTCCAGCTGACCGAGGAGAGAGCTCGTCGCCA 120
Db 140 GCGTCGATGTCGGGAGCGCCATCACGTCGCTGACGCTGACGAGAGAGAGCTCGTCGCCA 199
QY 121 CCATCGAGTACCTGTCGCGCTGACGAGGCGCCAGCAGCATGACCGTCCCGGGCGGCA 180
Db 200 CCATCGAATATCTGTCGCGCTGACGAGGCGTCAGCAGGAGAGAGCTCGTCGCCA 259
QY 181 CCGAGTCCCGGTTGAGACCGAGCAT 208
Db 260 TCGAGTCCCGTGGAAACCGAGCAT 287
RESULT 3
MTU318813
LOCUS
DEFINITION
MTU318813 616 bp DNA circular BCT 09-AUG-2002
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1763-97.
ACCESSION
AJ318813
VERSION
AJ318813.1 GI:22208402
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
Mycobacterium tuberculosis.
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1
Herrera, L., Jimenez, M.S. and Saez, J.A.
AUTHORS
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
REFERENCE
2 (bases 1 to 616)
Herrera, L.
AUTHORS
Direct Submission
TITLE
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
JOURNAL
FEATURES
source
1..616
/organism="Mycobacterium tuberculosis"
/isolate="1415-97"
/db_xref="taxon:1773"
gene
1..610
/genes="rpoB"
<1..>610
/genes="rpoB"
CDS
1..610
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87035.1"
/db_xref="GI:22208413"
/translation="LDIYRKLPGEPPTKESAOITLLENLFFKEKRYDLARVGRYKVNK
KLGHLVGEPIITSLTEDVDVATIEYLRHLEGQTTMTVPGGVEVPVETDDIDHFGNR
RLRTVGELIQNOIRVGMSRMERVMRTQDVEAITPQTLINIRPVAAIKEFFGTS
QPSQFMQNNPLSLGTHKRRLSALGPGGLSRERAGLEVRDHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match 80.0%; Score 166.4; DB 1; Length 610;
Best Local Similarity 87.5%; Pred. No. 1.5e-16;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGCGCGCTACAGGTCACAGAAGCTCG 60
Db 80 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGCGCGCTACAGGTCACAGAAGCTCG 139
QY 61 GCCTGCACCTGGCGATCCAGCTCCAGCTGACCGAGGAGAGAGCTCGTCGCCA 120
Db 140 GCGTCGATGTCGGGAGCGCCATCACGTCGCTGACGCTGACGAGAGAGAGCTCGTCGCCA 199
QY 121 CCATCGAGTACCTGTCGCGCTGACGAGGCGCCAGCAGCATGACCGTCCCGGGCGGCA 180
Db 200 CCATCGAATATCTGTCGCGCTGACGAGGCGTCAGCAGGAGAGAGCTCGTCGCCA 259
QY 181 CCGAGTCCCGGTTGAGACCGAGCAT 208
Db 260 TCGAGTCCCGTGGAAACCGAGCAT 287
RESULT 2
MTU318819
LOCUS
DEFINITION
MTU318819 610 bp DNA circular BCT 09-AUG-2002
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1417-97.
ACCESSION
AJ318819
VERSION
AJ318819.1 GI:22208414
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
Mycobacterium tuberculosis.
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1
Herrera, L., Jimenez, M.S. and Saez, J.A.
AUTHORS
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
REFERENCE
2 (bases 1 to 610)
Herrera, L.
AUTHORS
Direct Submission
TITLE
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
JOURNAL
FEATURES
source
1..610
/organism="Mycobacterium tuberculosis"
/isolate="1417-97"
/db_xref="taxon:1773"
gene
1..610
/genes="rpoB"
<1..>610
/genes="rpoB"
CDS
1..610
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87036.1"
/db_xref="GI:22208415"
/translation="LDIYRKLPGEPPTKESAOITLLENLFFKEKRYDLARVGRYKVNK
KLGHLVGEPIITSLTEDVDVATIEYLRHLEGQTTMTVPGGVEVPVETDDIDHFGNR
RLRTVGELIQNOIRVGMSRMERVMRTQDVEAITPQTLINIRPVAAIKEFFGTS
QPSQFMQNNPLSLGTHKRRLSALGPGGLSRERAGLEVRDHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match 80.0%; Score 166.4; DB 1; Length 610;
Best Local Similarity 87.5%; Pred. No. 1.5e-16;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGCGCGCTACAGGTCACAGAAGCTCG 60
Db 80 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGCGCGCTACAGGTCACAGAAGCTCG 139
QY 61 GCCTGCACCTGGCGATCCAGCTCCAGCTGACCGAGGAGAGAGCTCGTCGCCA 120
Db 140 GCGTCGATGTCGGGAGCGCCATCACGTCGCTGACGCTGACGAGAGAGAGCTCGTCGCCA 199
QY 121 CCATCGAGTACCTGTCGCGCTGACGAGGCGCCAGCAGCATGACCGTCCCGGGCGGCA 180
Db 200 CCATCGAATATCTGTCGCGCTGACGAGGCGTCAGCAGGAGAGAGCTCGTCGCCA 259
QY 181 CCGAGTCCCGGTTGAGACCGAGCAT 208
Db 260 TCGAGTCCCGTGGAAACCGAGCAT 287
```

gene /organism="Mycobacterium tuberculosis"
/isolate="1763-97"
/db_xref="taxon:1773"
1..615
/gene="rpoB"
CDS <1..>615
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87030.1"
/db_xref="GI:22208403"

translation="LDIVRKLRPGPPTKESAQTLLLENLFFKRYKRYDLARVGRYKVNK
KLGLHVGEPITSTIEDVVAIEYLRHGGOTTMTVPGGVEVPVETDDIDHFGNR
RLRTVGEILQNIORVGMSEMRVVRMTTQDVEAITPTQLINIRPVAAIKERFGTS
QLSIFDQNNPLSGLTHKRLSLGALGGSLRERAGLEVDRVHP"
BASE COUNT 125 a 191 c 201 g 99 t
ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 616;
Best Local Similarity 87.5%; Pred. No. 1.5e-16;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGCTACAAAGTCAACAAAGAGCTCG 60
|||||
Db 80 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGCTACAAAGTCAACAAAGAGCTCG 139
Qy 61 GCCTGCACGTCGGCGATCCGATCACCAGCTCCACGCTGACCGAGGAGAGCTCGTCGCCA 120
|||||
Db 140 GCCTGCATGTCGGCGAGCCCATCAGCTGCTGACGCTGACCGAGGAGAGCTCGTCGCCA 199
Qy 121 CCATCGAGTACCTGGTCCGCTCCAGGAGGCGGCACACGATGACCGTCCCGGGCGGCA 180
|||||
Db 200 CCATCGAATATCTGCTCCGCTTCACAGAGGTCAGACGATGACCGTCCCGGGCGGCG 259
Qy 181 CCGAGTGCCTGGTTGAGACCGAGCAT 208
Db 260 TCGAGTGCCTGGTGAACCGAGCAT 287

RESULT 4
MTU318815 618 bp DNA circular BCT 09-AUG-2002
LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
DEFINITION beta subunit, isolate 2540-97.
ACCESSION AJ318815
VERSION AJ318815.1 GI:22208406
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Mycobacterium;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE 1
AUTHORS Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
Isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 618)
AUTHORS Herrera,L.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN

FEATURES
source
1..618
/organism="Mycobacterium tuberculosis"
/isolate="2540-97"
/db_xref="taxon:1773"
gene 1..618
CDS <1..>618

/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87032.1"
/db_xref="GI:22208407"
translation="TDEALLDIYRKLRPGPPTKESAQTLLLENLFFKRYKRYDLARVGR
YKYNKLGHLHVGEPITSTIEDVVAIEYLRHGGOTTMTVPGGVEVPVETDDID
HFGNRLRTVGEILQNIORVGMSEMRVVRMTTQDVEAITPTQLINIRPVVAAIKE
FFGTQSUSQFMDQNNPLSGTHRRRLALPGGLSRERAGLEVDRV"
BASE COUNT 124 a 190 c 207 g 97 t
ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 618;
Best Local Similarity 87.5%; Pred. No. 1.5e-16;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGCTACAAAGTCAACAAAGAGCTCG 60
|||||
Db 95 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGCTACAAAGTCAACAAAGAGCTCG 154
Qy 61 GCCTGCACGTCGGCGATCCGATCACCAGCTCCACGCTGACCGAGGAGAGCTCGTCGCCA 120
|||||
Db 155 GCCTGCATGTCGGCGAGCCCATCAGCTGCTGACGCTGACCGAGGAGAGCTCGTCGCCA 214
Qy 121 CCATCGAGTACCTGGTCCGCTCCAGGAGGCGGCACACGATGACCGTCCCGGGCGGCA 180
|||||
Db 215 CCATCGAATATCTGCTCCGCTTCACGAGGTCAGACCGATGACCGTCCCGGGCGGCG 274
Qy 181 CCGAGTGCCTGGTTGAGACCGAGCAT 208
Db 275 TCGAGTGCCTGGTGAACCGAGCAT 302

RESULT 5
MTU318817 618 bp DNA circular BCT 09-AUG-2002
LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
DEFINITION beta subunit, isolate 2348-98.
ACCESSION AJ318817
VERSION AJ318817.1 GI:22208410
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis
Bacteria; Actinobacteria; Actinobacteridae; Mycobacterium;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE 1
AUTHORS Herrera,L., Jimenez,M.S. and Saez,J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 618)
AUTHORS Herrera,L.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN

FEATURES
source
1..618
/organism="Mycobacterium tuberculosis"
/isolate="2348-98"
/db_xref="taxon:1773"
gene 1..618
CDS <1..>618
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87034.1"
/db_xref="GI:22208411"

/translation="TDEALLDIYKLRPGEPPTKESAQTLLNLFKEKRYDLRVGR
YKVNKKLGLHVGEPTSSLTEDVVAITYEYLVRLHEGQTTMTVPGGVEVPVETDDID
HFGNRLRTVGLGELIQIRVGMSEMRVVRMTTQDVEAITPOTLINIRPVVAIKE
FPGTSQSPFMGQNNPLSLTHKRLSALPGGLSRERAGLEVRDV"
BASE COUNT 124 a 192 c 207 g 95 t
ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 618;
Best Local Similarity 87.5%; Pred. No. 1.5e-16;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTGCCCGGTAGGCGCTACAAGGTCAACAAGAACTCG 60
Db TCAAGGAGAGCGCTACGACCTGCCCGGTAGGCGCTACAAGGTCAACAAGAACTCG 60
QY 95 TCAAGGAGAGCGCTACGACCTGCCCGGTAGGCGCTACAAGGTCAACAAGAACTCG 154
Db TCAAGGAGAGCGCTACGACCTGCCCGGTAGGCGCTACAAGGTCAACAAGAACTCG 154
QY 61 GCCTGCAGCTCGGGATCCCATCACAGTCCACGCTGACCGAGGAGAGAGCTGTCGCCA 120
Db TCGTGCATGTCGGCGAGCCCATCAGCTGCTGACCGTACCGGAGGAGAGAGCTGTCGCCA 214
QY 121 CCATCGAGTACCTGCTCGCGCTGCACGAGGCGCCACACAGCATGACCGTCCCGGGCGGCA 180
Db CCATCGAATATCTGCTCCGCTTCACGAGGCTCAGACGAGGTTCACCGAGGCTCGCGGGCGG 274
QY 181 CCGAGTGCCTGCTGACACGACGACAT 208
Db TCGAGTGCCTGCTGACACGACGACAT 302

RESULT 6
MTU318814
LOCUS
DEFINITION
MTU318814 633 bp DNA circular BCT 09-AUG-2002
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1058-97.
ACCESSION
AJ318814
VERSION
AJ318814.1 GI:22208404
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE
1
Herrera, L., Jimenez, M. S. and Saez, J. A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
Unpublished
JOURNAL
2 (bases 1 to 633)
Herrera, L.
Direct Submission
AUTHORS
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN
JOURNAL

FEATURES
source
1. .633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
gene
1. .633
/gene="rpoB"
CDS
<1..>633
/gene="rpoB"
/codon_start=1
/product="RNA polymerase beta subunit"
/protein_id="CAC87031.1"
/db_xref="GI:22208405"

/translation="DNTVGTDEALLDIYKLRPGESPPTKESAQTLLNLFKEKRYDL
ARVGRYKVNKKLGLHVGEPTSSLTEDVVAITYEYLVRLHEGQTTMTVPGGVEVPVE
TDDIDHFGNRLRTVGLGELIQIRVGMSEMRVVRMTTQDVEAITPOTLINIRPVV
AAIKEFFTSQSPFMGQNNPLSLGTYKRLSALPGGLSRERAGLEVRDV"
BASE COUNT 129 a 195 c 210 g 99 t
ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 633;
Best Local Similarity 87.5%; Pred. No. 1.5e-16;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTGCCCGGTAGGCGCTACAAGGTCAACAAGAACTCG 60
Db TCAAGGAGAGCGCTACGACCTGCCCGGTAGGCGCTACAAGGTCAACAAGAACTCG 169
QY 61 GCCTGCAGCTCGGGATCCCATCACAGTCCACGCTCACCAGGAGGAGAGCTGTCGCCA 120
Db TCGTGCATGTCGGCGAGCCCATCAGCTGCTGACCGTACCGGAGGAGAGAGCTGTCGCCA 229
QY 121 CCATCGAGTACCTGCTCGCGCTGCACGAGGCGCCACACAGCATGACCGTCCCGGGCGGCA 180
Db CCATCGAATATCTGCTCCGCTTGCACGAGGCTCAGACCAACGATGACCGTTCGGGGCGGCG 289
QY 181 CCGAGTGCCTGCTGAGACGACGACAT 208
Db TCGAGTGCCTGCTGAGACGACGACAT 317

RESULT 7
MTU318816
LOCUS
DEFINITION
MTU318816 637 bp DNA circular BCT 09-AUG-2002
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1255-98.
ACCESSION
AJ318816
VERSION
AJ318816.1 GI:22208408
KEYWORDS
RNA polymerase beta subunit; rpoB gene.
SOURCE
Mycobacterium tuberculosis.
ORGANISM
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.

REFERENCE
1
Herrera, L., Jimenez, M. S. and Saez, J. A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
Unpublished
JOURNAL
2 (bases 1 to 637)
Herrera, L.
Direct Submission
AUTHORS
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN
JOURNAL

FEATURES
source
1. .637
/organism="Mycobacterium tuberculosis"
/isolate="1255-98"
/db_xref="taxon:1773"
gene
1. .637
/gene="rpoB"
CDS
<1..>637
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87033.1"
/db_xref="GI:22208409"

/translation="VGTDEALLDIYKLRPGEPPTKESAQTLLNLFKEKRYDLARV
GRYKVNKKLGLHVGEPTSSLTEDVVAITYEYLVRLHEGQTTMTVPGGVEVPVETDD
IDHFGNRLRTVGLGELIQIRVGMSEMRVVRMTTQDVEAITPOTLINIRPVVAAI
KEFFTSQSPFMGQNNPLSKRMRMFGGLSRERAGLEVRDVHPSH"
BASE COUNT 128 a 198 c 210 g 101 t
ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 637;
Best Local Similarity 87.5%; Pred. No. 1.5e-16;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTGCCCGGTAGGCGCTACAAGGTCAACAAGAACTCG 60

```

|||||
Db 101 TCAAGGAGAGCGCTACGACCTGGCCCGCTCGGTGCTATTAAGTCAACAGAAAGCTCG 160
QY 61 GCCTGCACGTGCGGATCCGATACACAGCTCCACGCTGACCGAGGAAGACGTCGTGCGCA 120
Db 161 GGCTGCATGTGCGGAGCGCATCACCTCGCTCGACGCTACCGAAGAAGACGTCGTGGCCA 220
QY 121 CCATCGAGTACCTGGTCCGCTGACGAGGCGCAGACACGATGACCGTCCCGGGCGGCA 180
Db 221 CCATCGAATATCTGGTCCGCTGACGAGGGTCAGACCACGATGACCGTTCCGGGCGGCG 280
QY 181 CCGAGGTGCGCGTTGAGACCGACGACAT 208
Db 281 TCGAGGTGCGCGTGGAAACCGAGGACAT 308

RESULT 8
MTU318821
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
FEATURES
Source
gene
CDS

MTU318821
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1071-98.
AJ318821
RNA polymerase beta subunit; rpoB gene.
Mycobacterium tuberculosis.
Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
Unpublished
2 (bases 1 to 639)
Herrera, L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
Location/Qualifiers
1..639
/organism="Mycobacterium tuberculosis"
/isolate="1071-98"
/db_xref="taxon:1773"
1..639
/gene="rpoB"
<1..>639
/gene="rpoB"
/codon_start=1
/product="RNA polymerase beta subunit"
/db_xref="GI:22208419"
/translation="ARTDEALLDYRKLRCPEPTKESQTLLENLFKRYDLARY
GRYKNKGLHVGPISTLTEDVATIEYLVRHGGTTMTVPGVBPVETDD
IDHFNKRLRTVGLIQNOIRVMSRMVVRRTQDVEAITPQTLINRPAAAI
KEFFTQSLSQFMVQNNPLSLGTQKRLSALPGSLSRERAGLEVDRVDPHS"

BASE COUNT 126 a 202 c 212 g 99 t
ORIGIN

Query Match 80.0%; Score 166.4; DB 1; Length 639;
Best Local Similarity 87.5%; Pred. No. 1.5e-16;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGGCTACAGGTCAACAGAAAGCTCG 60
Db 101 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGGCTATTAAGTCAACAGAAAGCTCG 160
QY 61 GCCTGCACGTGCGGATCCGATACACAGCTCCACGCTGACCGAGGAAGACGTCGTGCGCA 120
Db 161 GGCTGCATGTGCGGAGCGCATCACCTCGCTCGACGCTACCGAAGAAGACGTCGTGGCCA 220
QY 121 CCATCGAGTACCTGGTCCGCTGACGAGGCGCAGACACGATGACCGTCCCGGGCGGCA 180
Db 221 CCATCGAATATCTGGTCCGCTGACGAGGGTCAGACCACGATGACCGTTCCGGGCGGCG 280
QY 181 CCGAGGTGCGCGTTGAGACCGACGACAT 208
Db 281 TCGAGGTGCGCGTGGAAACCGAGGACAT 308

RESULT 9
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT 182 a 302 c 330 g 156 t
ORIGIN

Query Match 80.0%; Score 166.4; DB 6; Length 970;
Best Local Similarity 87.5%; Pred. No. 1.4e-16;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTAGACCTGGCCCGGTAGCGGCTACAGGTCAACAGAAAGCTCG 60
Db 26 TCAAGGAGAGCGCTAGACCTGGCCCGGTAGCGGCTATTAAGTCAACAGAAAGCTCG 85
QY 61 GCCTGCACGTGCGGATCCGATCACGATCACGCTCCACGCTGACCGAGGAAGACGTCGTGCGCA 120
Db 86 GCCTGCATGTGCGGAGCGCCATCACGCTCGTCGACGCGTGACCGAAGAAGACGTCGTGGCCA 145
QY 121 CCATCGAGTACCTGGTCCGCTGACGAGGCGCAGACACGATGACCGTCCCGGGCGGCA 180
Db 146 CCATCGAATATCTGTGCTCGCTTGACGAGGTCAGACGATGACCGTCCGGGCGGCG 205
QY 181 CCGAGGTGCGCGTTGAGACCGACGACAT 208
Db 206 TCGAGGTGCGCGTGGAAACCGAGGACAT 233

RESULT 10
AX111339
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
FEATURES

AX111339
Sequence 2072 from Patent WO0123604.
AX111339
AX111339.1 GI:13927631
Mycobacterium tuberculosis.
Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 3534)
Bergeron, M.G., Boissinot, M., Huletsky, A., m Nard, C., Ouellette, M.,
Picard, F.J. and Roy, P.H.
Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
Patent: WO 0123604-A 2072 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
Location/Qualifiers

```

```

QY 121 CCATCGAGTACCTGGTCCGCTGACGAGGCGCAGACAGATGACCGTCCCGGGCGGCA 180
|||||
Db 221 CCATCGAATATCTGTGCTCGCTTGACGAGGTCAGACGATGACCGTCCGGGCGGCG 280
QY 181 CCGAGGTGCGCGTTGAGACCGACGACAT 208
|||||
Db 281 TCGAGGTGCGCGTGGAAACCGAGGACAT 308

RESULT 9
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
FEATURES
source
BASE COUNT 182 a 302 c 330 g 156 t
ORIGIN

Query Match 80.0%; Score 166.4; DB 6; Length 970;
Best Local Similarity 87.5%; Pred. No. 1.4e-16;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTAGACCTGGCCCGGTAGCGGCTACAGGTCAACAGAAAGCTCG 60
Db 26 TCAAGGAGAGCGCTAGACCTGGCCCGGTAGCGGCTATTAAGTCAACAGAAAGCTCG 85
QY 61 GCCTGCACGTGCGGATCCGATCACGATCACGCTCCACGCTGACCGAGGAAGACGTCGTGCGCA 120
Db 86 GCCTGCATGTGCGGAGCGCCATCACGCTCGTCGACGCGTGACCGAAGAAGACGTCGTGGCCA 145
QY 121 CCATCGAGTACCTGGTCCGCTGACGAGGCGCAGACACGATGACCGTCCCGGGCGGCA 180
Db 146 CCATCGAATATCTGTGCTCGCTTGACGAGGTCAGACGATGACCGTCCGGGCGGCG 205
QY 181 CCGAGGTGCGCGTTGAGACCGACGACAT 208
Db 206 TCGAGGTGCGCGTGGAAACCGAGGACAT 233

RESULT 10
AX111339
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
FEATURES

AX111339
Sequence 2072 from Patent WO0123604.
AX111339
AX111339.1 GI:13927631
Mycobacterium tuberculosis.
Mycobacterium tuberculosis.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 3534)
Bergeron, M.G., Boissinot, M., Huletsky, A., m Nard, C., Ouellette, M.,
Picard, F.J. and Roy, P.H.
Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
Patent: WO 0123604-A 2072 05-APR-2001;
Infectio Diagnostic (I.D.I.) INC. (CA)
Location/Qualifiers

```

source	1. .3534		/organism="Mycobacterium tuberculosis"		/strain="Rv"		/db_xref="taxon:1773"		BASE COUNT		679 a		1081 c		1188 g		586 t		ORIGIN	
Query Match	80.0%;		Score 166.4;		DB 6;		Length 3534;													
Best Local Similarity	87.5%;		Pred. No. 1.1e-16;																	
Matches 182;	Conservative 0;		Mismatches 26;		Indels 0;		Gaps 0;													
QY	1		TCAAGGAGAGCGCTAGCAGCTGCCCGGTAGCCGCTACAGGCTCAACAAGAGCTCG		60															
Db	902		TCAAGGAGAGCGCTAGCAGCTGCCCGGTAGCCGCTACAGGCTCAACAAGAGCTCG		961															
QY	61		GCCTGCAGCTGCGGATCCGATCACCAGCTCCAGCTCCAGGAGAGAGCTCGTCGCCCA		120															
Db	962		GCCTGCATGTCCGGAGGCCATCAGCTCGCTGCGCTGACCGAAGAGAGCTCGTGCCCA		1021															
QY	121		CCATCGAGTACTGTGTCGCTGACGAGGCCAGCAGCATGACCGTCCCGGGCGGCA		180															
Db	1022		CCATCGAATATCTGTGTCGCTGACGAGGGTCAGACCATGACCGTTCCTCGGGCGCG		1081															
QY	181		CCGAGGTCCCGTTGAGACCGACGACAT		208															
Db	1082		TCGAGGTCCCGTGGAAACCGACGACAT		1109															
RESULT 11																				
MTU12205																				
LOCUS																				
DEFINITION																				
ACCESSION																				
VERSION																				
KEYWORDS																				
SOURCE																				
ORGANISM																				
REFERENCE																				
AUTHORS																				
TITLE																				
JOURNAL																				
REFERENCE																				
AUTHORS																				
TITLE																				
JOURNAL																				
FEATURES																				
source																				
gene																				
CDS																				

BASE COUNT		723 a	1173 c	1293 g	664 t	ORIGIN
Query Match		80.0%; Score 166.4; DB 1; Length 3853;				
Best Local Similarity		87.5%; Pred. No. 1e-16;				
Matches 182; Conservative		0;	Mismatches 26;	Indels 0;	Gaps 0;	
QY	1	TCAAGGAGAGCGCTACGACCTGCCCGGTAGCCGCTACAAAGGTCAACAAGAGCTCG 60				
Db	1477	TCAAGGAGAGCGCTAGCAGCTGCCCGGTGCTATAGGTCACACAGAGCTCG 1536				
QY	61	GCCTGCAGCTGCGCGATCCGATCACCAGCTCCAGCTCCAGGAGAGAGCTCGTCGCCCA 120				
Db	1537	GGCTGCATGTCCGGAGGCCATCAGCTCGCTGCGTATAGGTCACACAGAGCTCGTCGCCCA 1596				
QY	121	CCATCAGTACCTGGTCCGCTCCAGGAGGCCAGCAGATGACCGTCCCGGGCGGCA 180				
Db	1597	CCATCGAATATCTGGTCCGCTTGCAGGAGGTGACACCATGACCGTTCCTCGGGCGGCG 1656				
QY	181	CCGAGGTCCCGTTGAGACCGACGACAT 208				
Db	1657	TCGAGGTCCCGTGGAAACCGACGACAT 1684				
RESULT 12						
MSGRPOB 5084 bp DNA linear BCT 13-SEP-1994						
LOCUS Mycobacterium tuberculosis RNA polymerase beta-subunit (rpoB)						
DEFINITION gene, complete cds and RNA polymerase beta'-subunit rpoC gene, partial cds.						
ACCESSION L27989						
VERSION L27989.1 GI:468333						
KEYWORDS RNA polymerase beta-subunit; rpoB gene.						
SOURCE Mycobacterium tuberculosis (strain Rv) DNA.						
ORGANISM Mycobacterium tuberculosis						
REFERENCE Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.						
1 (bases 1 to 5084)						
MILLER, L.P., CRAWFORD, J.T. and SHINNICK, T.M.						
TITLE The rpoB gene of Mycobacterium tuberculosis						
JOURNAL Antimicrob. Agents Chemother. 38 (4), 805-811 (1994)						
MEDLINE 94304130						
PUBMED 8031050						
FEATURES						
Location/Qualifiers						
1. .5084						
/organism="Mycobacterium tuberculosis"						
/strain="Rv"						
/db_xref="taxon:1773"						
1065..4598						
/gene="rpoB"						
1065..4598						
/gene="rpoB"						
/codon_start=1						
/evidence=experimental						
/transl_table=11						
/product="RNA polymerase beta-subunit"						
/protein_id="AAA21416.1"						
/db_xref="GI:468334"						
/translation="MLEGCILADRSQKTAASPSRPOSSNNNSVPGAPNRVSAFKL						

REPLEVPGLLDVQTSFEWLGSPRWRESAERGDVNPVGGLEEVLYELSLPIEDFSGS
MSLSFSDREFDDVKAPVDECKDMTYAAPLFVTAERINNNTGEIKSOTVPMGDFPMN
TEKGFTINGERRVVSQVSRPGVYFDETDKSDKTLHSVKVLPBSRGANLEDDVK
RTVGVRIDRRKRPQVTVLLKALGWTSEQIVVERGFSIMRSTLEKONTVGTDEALLD
IYRKLRPEPTKSAQTLLENLFFKERYDLARVGRYKVKKLGHLHVGEPITSTLT
EDDVATIEYLVRHEGQTTMTVPGVPEVPTDDIDHFGNRLRTVGLHQNQIRVVG
MSRMERVVERMTTODVEAITPQTLINIRPVAAIKFEGTSQLSQFMDQNNPLSGLT
HKRRLSALCPGSLSRERAGLEVRDVHPSHYGRMCPETPEGNIGLIGLSVYARVNP
FCFIETPVKVVGVDSDEIYVLTADAEDRHVVAQANSPIDADGDFEVPVLYVRKAG
EVEYVPSSEYDMVSPRQMSVATAMIPFLEHDDANRALGANQORQAVPLVSEAP
LVGTGMEURAAIDAATSSQESGVIEVSADYITVMHNGTTRRTYMRKFAHSNHTC
ANQCPITDAGDRAGVQIADGCTDDGEMALGNLLVAIMPWECHNYDELLINRL
VEDVLGTSIHIEHEIDARTKLGAEITRDPINISDEVLDADLDEGRGIVRGAVERDG
DILVKGTPKGETELTPEERLLRAIFGEKAREVDTSLKVPHGSGKVGIGIRVFSRED
EDELPAQYNELVRVVAQKRKISDGDKLAGRHNKGVIKILPVEDMFPADGTPVDI
ILNTHGVPFRMNIQIILETHLGCAHSGWKVDAAKGVDPDAAARLPDELEERAHNAIVS
TVFQGAQEAELQGLSCTLPNRDGDVLVDADGKAMLFDRGSGEPFPVPTVGYMYIM
KLHLVDKIHARSTGYSMTIQPLGKAQFGQFGGEMECWAMQAYGAAYTLQEL
TKISDQTVGRVYEAIVKGENIPGPIPESGFKVLLKELQSLCLNNEVLSSDGAAIL
REGEDEDLERAAANLGINLSRNESASFEDLA"

4641.. 5084
/gene="rpoC"
4641..>5084
/gene="rpoC"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta'-subunit"
/protein_id="AA21417.1"
/db_xref="GI:537608"
/translation="MLDVNFFDELRIGLATAEDIRQWSYGEYKPKPETINYRTLKPEKD
GLFCEKIFGTPDMWECYCGKTKYRFRFGIICERCGVEVTRAKYRRMRHGHTELAAPVT
HIWFKGVPSRLGYLLDLAPKDLKLEIYFAAYVITSVDEEMRHNEL"

BASE COUNT 969 a 1534 c 1691 g 890 t

gene
CDS

Query Match 80.0%; Score 166.4; DB 1; Length 5084;
Best Local Similarity 87.5%; Pred. No. 9.9e-17;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1 TCAAGGAGAGCGCTAGCAGCTGCCCGGCTAGCGCTACAGGTCACAGAAGCTCG 60
|||||
Db 1966 TCAAGGAGAGCGCTAGCAGCTGCCCGGCTCGGCTATAGAGTCAACAAGAAGCTCG 2025
|||||
QY 61 GCCTCGACGTCGGCGATCCCATCAGGCTCCAGCGTGACCGAGGAGAGCGTCTCGTCCGCA 120
|||||
Db 2026 GCCTGATGTCGGGAGCCCATCAGCTGCTGACGCTGACCGAAGAGCGTCTGCGCCA 2085
|||||
QY 121 CCATCGAGTACCTGTCGCTGACAGAGGCGGACGACGATGACCGTCCCGGGCGGCA 180
|||||
Db 2086 CCATCGAATATCTGTCGCTGACAGAGGTCAGACGATGACCGTTCGCGGCGGCG 2145
|||||
QY 181 CCGAGGTGCGCGTTGAGACCGAGCAT 208
|||||
Db 2146 TCGAGGTGCGCGGTGGAACCGAGCAT 2173
|||||

RESULT 13
LOCUS AE006964 19352 bp DNA linear BCT 27-APR-2001
DEFINITION Mycobacterium tuberculosis CDC1551, section 50 of 280 of the complete genome.
ACCESSION AE006964 AE000516
VERSION AE006964.1 GI:13880217
KEYWORDS
SOURCE Mycobacterium tuberculosis CDC1551.
ORGANISM Mycobacterium tuberculosis CDC1551
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 19352)
REFERENCE
AUTHORS Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M.,

Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains
Unpublished
2 (bases 1 to 19352)
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A. and Bishai, W.
Direct Submission
Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
Location/Qualifiers
1..19352
/organism="Mycobacterium tuberculosis CDC1551"
/strain="CDC1551"
/db_xref="taxon:83331"
/note="clinical strain"
163..3699
/gene="rpo695"
163..3699
/gene="rpo695"
/note="similar to GB:L27989 GB:L05910 GB:U12205 SP:P47766 PID:149992; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA-directed RNA polymerase, beta subunit"
/protein_id="AAK44921.1"
/db_xref="GI:13880218"
/translation="MLESCILADRSQSKTAASPSRPOSSNNVSPGAPNRYSFAKL
REPLEVPGLLDVQTSFEWLGSPRWRESAERGDVNPVGGLEEVLYELSLPIEDFSGS
MSLSFSDREFDDVKAPVDECKDMTYAAPLFVTAERINNNTGEIKSOTVPMGDFPMN
TEKGFTINGERRVVSQVSRPGVYFDETDKSDKTLHSVKVLPBSRGANLEDDVK
RTVGVRIDRRKRPQVTVLLKALGWTSEQIVVERGFSIMRSTLEKONTVGTDEALLD
IYRKLRPEPTKSAQTLLENLFFKERYDLARVGRYKVKKLGHLHVGEPITSTLT
EDDVATIEYLVRHEGQTTMTVPGVPEVPTDDIDHFGNRLRTVGLHQNQIRVVG
MSRMERVVERMTTODVEAITPQTLINIRPVAAIKFEGTSQLSQFMDQNNPLSGLT
HKRRLSALCPGSLSRERAGLEVRDVHPSHYGRMCPETPEGNIGLIGLSVYARVNP
FCFIETPVKVVGVDSDEIYVLTADAEDRHVVAQANSPIDADGDFEVPVLYVRKAG
EVEYVPSSEYDMVSPRQMSVATAMIPFLEHDDANRALGANQORQAVPLVSEAP
LVGTGMEURAAIDAATSSQESGVIEVSADYITVMHNGTTRRTYMRKFAHSNHTC
ANQCPITDAGDRAGVQIADGCTDDGEMALGNLLVAIMPWECHNYDELLINRL
VEDVLGTSIHIEHEIDARTKLGAEITRDPINISDEVLDADLDEGRGIVRGAVERDG
DILVKGTPKGETELTPEERLLRAIFGEKAREVDTSLKVPHGSGKVGIGIRVFSRED
EDELPAQYNELVRVVAQKRKISDGDKLAGRHNKGVIKILPVEDMFPADGTPVDI
ILNTHGVPFRMNIQIILETHLGCAHSGWKVDAAKGVDPDAAARLPDELEERAHNAIVS
TVFQGAQEAELQGLSCTLPNRDGDVLVDADGKAMLFDRGSGEPFPVPTVGYMYIM
KLHLVDKIHARSTGYSMTIQPLGKAQFGQFGGEMECWAMQAYGAAYTLQEL
TKISDQTVGRVYEAIVKGENIPGPIPESGFKVLLKELQSLCLNNEVLSSDGAAIL
REGEDEDLERAAANLGINLSRNESASFEDLA"

3744.. 7694
/gene="MT0696"
3744.. 7694
/gene="MT0696"
/note="similar to SP:P37871; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA-directed RNA polymerase, beta-prime subunit"
/protein_id="AAK44922.1"
/db_xref="GI:13880219"
/translation="MLDVNFFDELRIGLATAEDIRQWSYGEYKPKPETINYRTLKPEKD
GLFCEKIFGTPDMWECYCGKTKYRFRFGIICERCGVEVTRAKYRRMRHGHTELAAPVT
HIWFKGVPSRLGYLLDLAPKDLKLEIYFAAYVITSVDEEMRHNELSTLEAEMAVERK
AVEDQRDELEAQAQLEADLALEAKADARRYKVDGGERMRGIRORAEARELDR
LEDIWSFTTKAPKOLVDENLYRELVDYGEYFTGAMGAEIOKLTENFDIDAEAS
LRDVRNGKGOKKLRLKLVAAFOQSGNSPMGVLDAPVPIPELRPMVOLDGGR
FATSDNDLRRVINRNRLKRLIDLGAPEIYNNKRMLOESYDALFDNGRRRPVT
GPNRPLKUSDLKQGRFRQNLKGRVDYSGRVVVGPOKLHQCGGLKMLALE
LFKPFVKRLVDLNLHAQNIKSRMVERQPVQWVLEEVIAEHPVLNLRAPTLHRLG
IQAFEPMLVEGKATQLHPLVCEAFNADFDGDMAVHLPLSAEAQAERILMLSSNNIL

```
SPASGRPLAMPRLDMVTGLVYLTTEVPEDTGEYQOPASGDHPETGVYSSPAEAIMAADR
GVLVRAKIVRLTQLRPPVEIEAELEFGHSGQOPGDWAEITLGRVMENELPLGYP
FVNQKHKKVQAAIINDLAERYPMIYVQATVQKLDAGFYWATRSVTVSMADVLVPP
RKEILDHYEERADKVEKQFQALGNHDERNEALVEIKWEATEDEVGQALREHYPDNP
IITIVSDGNTQFQKTRTAGMKGLVTPKGEFIPRPVKSSEFTEGQALREHYPDNP
RKLADALTADSGYLTRLRLVDVSDQVIVREHDCOTERGIIVVELAERAPDGLIRDP
YIETSAVARTLGTDAVDEAGNVIVERQDQGOPEIDALLAAGITOVKVRSLVTCATST
GVATCYGRSMATGKLVLDIGEAGVIVAAGSIEGPGTQLTMRFTHOQGVSDITGGLPR
VQELFEARPRGKAPTADVTGRVRLDGERFYKITIVPDDGGEVGYDKISRQRLRV
FKHEDSERVSDGHHVEVGOQLMEGSDAPHEVLRVQGGPREVQIHLVRVQEVYRAQG
VSLHDKHIEVILQMLRRVITIDSGSTFLPGLSLIDRDAEFENRRRVAAEGGEPAAGR
PVMHGTHKSLATLWLSAASFQETTRVLTRVLDAAINCDSKLNKLENVILIGKLIPAGT
GINRYNIAVQPTTEARAAYTIPSYEDQYISPDFGAATCAAVPLDDYGSYDR"
/genes="MT0697"
complement(7691..8065)
/genes="MT0697"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK4923.1"
/db_xref="GI:13880220"
/transl_table=11
/translation="MFDSSAAITNPGHAWASAMERSGLLECVAAGLDEQPFGEFTADKL
NPDGSSRRVPRRQADGIGTATHVERGGQSGQAGVVPQRMHGFPPALAMQDRLIHH
GEQTQNRIAQAFVRFCVCSPT"
complement(8058..9572)
/genes="MT0698"
/notes="This region contains an authentic frame shift and
is not the result of a sequencing artifact; identified by
Glimmer2; putative; conserved hypothetical protein,
authentic frameshift"
10167..10925
/genes="MT0699"
10167..10925
/genes="MT0699"
/notes="identified by match to PFAM protein family HMM
PF01261"
/codon_start=1
/transl_table=11
/product="AP endonuclease, family 2"
/protein_id="AAK4924.1"
/db_xref="GI:13880221"
/transl_table=11
/translation="MLIGSHVSTPDLAAAEAGADVQVIFLGNPQSKAPKPRDDAA
ALKAATLPIVYHAPYILNLSANNRVRIPSRLILOETCAAADIGAAAVIVHGHHVAD
DNIDKGFQWRKALDRLETPVYLVNTAGNDHANAARFDTIARLWDVIGDTGIGFC
LDYCHTWAGEALTDAVDRIKAITGRIDLHVHCNDSRDEAGSGRDRHANLGSQIDPDL
LVAAVRAAGAPVICETADQGRKDDIAFLRITGS"
10957..11799
/genes="MT0700"
10957..11799
/genes="MT0700"
/notes="similar to GB:U00012 PID:466863; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="hydrolase/esterase, putative"
/protein_id="AAK4925.1"
/db_xref="GI:13880222"
/transl_table=11
/translation="MLRRVAILLAVALAFAGSGGTRLAAGFGNGNSVHTLVDGAGR
SYRLKPVGLPSSAPLVMHLHGFGSAKARSYGWDELASEKELVAYPDGYHRAWN
ANGKCGGREGVDDIGFVRVAVADIANNYSIDPARVYVTGMSGNGAIMSYTLACNT
SIFAAIYVSGTQLDPCQSPRPVSVIHIGTADPLVRYHGGPGAGFARIDGPPVPLN
AFWRVNRGALDITTEGPTTSGATCADNRRRVLLTVDDAGHRWPSFATQTLWRFFA
AHRF"
11859..13487
/genes="MT0701"
11859..13487
/genes="MT0701"
/notes="similar to SP:P33224 GB:L20915 PID:457172
PID:457174 PID:537028; identified by sequence similarity;
putative"
/codon_start=1
```

```
/transl_table=11
/product="acyl-CoA dehydrogenase, putative"
/protein_id="AAK4926.1"
/db_xref="GI:13880223"
/transl_table=11
/translation="MSDTHVTVNQVPPLENTNPASSPVLIEALIQEGQMGLEDEVNEV
GAISASCOQWGLADNRNRIILTHDAYGYRDEVEYDPAYHELMRTAITHGHHAAP
WADDPGAAHVRAAKTSVMTVEPGHICPISMTYAVVPALRYNSLAAYEPPLTSREY
DPELKPATTKAGITAGNSMTSEKQSGDVRACTTQATPNADGSYSLTGKHKWTSAPMCD
IFLVLAQAPDSCGLFLLPRVLPDGTGRNRMFLQRLDKLGNHANASSEVEYDGAVALV
GEEGVPTIITEMVNLRLDLCALGSMTGLTRAVHHAQHRKAFAYLIDQPLMRN
VLADIAVEAEATIVAMRMAGATDNVAGNETEALLRRIGLAAAKYVWCKRSTAAHAE
ALEICGGNGYVEDSGMRLYREAPLMGWEGSGNVSDTLRAMATRPACVEVLFDEL
ARSAGQDRLDGHVERLRLPOLGDLDTIGYRARKTAEDICLALQGLSLVLRHGHGPAVAEA
FLATPLGGQGGAYGTMPLAGLDLAPILRALVKG"
13498..14436
/genes="MT0702"
13498..14436
/genes="MT0702"
/notes="similar to GP:3885480; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="enoyl-CoA hydratase/isomerase family protein"
/protein_id="AAK4927.1"
/db_xref="GI:13880224"
/transl_table=11
/translation="MTHAIRPVDFDNLKMTYEVTVGTRIARTFNRPKGNAILIADTPL
ELSALVERADLDPGVHVILVSGREGGFCAGFDLSAYAGSSSTGGGAYGQTVLDGKT
QAVNHLPNQPMIDYQOMSRFVRGASFALMHADKPTVYKIHGYCVAGGTDLALHADO
YAAADAKIGYPTTRVGVPAAGLWHRFLDQRAKRLITFTGDCITGAQAAEMGLAVEA
PEPADLDERLVARIAALPNQILIMVKLALNSALLQOQVATSRMSTVFDGAARHT
PEGHAFVADAVEHGFDAVRRRDEPFQDYGQASRV"
14439..15161
/genes="MT0703"
14439..15161
/genes="MT0703"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
Query Match 80.0%; Score 166.4; DB 1; Length 19352;
Best Local Similarity 87.5%; Pred. No. 7.7e-17;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGGTACCAAGCTCAACAAGAGCTCG 60
|||||
Db 1064 TCAAGGAGAGCGCTACGACCTGGCCCGGTAGCGGTATTAAGGTCAACAAGAGCTCG 1123
|||||
Qy 61 GCCTGCACGTCGCGCATCCGATCACCAGCTCCACGAGGAGGAGCGTCGCGCCA 120
|||||
Db 1124 GCCTGCATGTCGGGAGCGCATCAGCTCGTCCAGCGTACCAGGAGGAGCTCG 1183
|||||
Qy 121 CCATCAGTACCTGTCGCGCTGACGAGGCGGACACATGATGACCGCGCGGCGCA 180
|||||
Db 1184 CCATCGAATATCTGTCGCGCTTGCACGAGGCTCAGACCATGATGACCGTCCGCGCGCG 1243
|||||
Qy 181 CCGAGTGCCGCTTGAGACCGGACGACAT 208
|||||
Db 1244 TCAGGTGCGCGTGGAAACCGGACGACAT 1271
|||||
RESULT 14
MTC1376
LOCUS
DEFINITION
ACCESSION 295972 AL123456
VERSION 295972.1 GI:3261790
KEYWORDS
SOURCE
ORGANISM
Mycobacterium tuberculosis H37Rv.
Mycobacterium tuberculosis H37Rv
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
```


probablemembrane protein, similar to YPRB_ECOLI_P33774
hypothetical24.3 kd protein (unf 1) (217 aa), fasta
scores: opt: 174 z-score: 215.8 E(): 5.3e-05, (25.6%
identity in 223 aa overlap). Also similar to MTCY359.10
(28.7% identity in 178 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0658c"
/protein_id="CAB09376.1"
/db_xref="GI:2143303"
/db_xref="SPTREMBL:O06781"
/translation="MEAGRADTVAPSHRWGLGAFVLVFLVASTSLAVVLTHGVP
SAGVLALAAPTVVVAGLAIIITRLNGRLTDLRLRWSHRLGLRMFGFGMLVT
IPASLVYTAIVGPENASAVRIFGGRASWPALVFLVVFVAPLCEEEIIRGLLWG
AVDRWGMALVTVTVFVFAHLEFARAPLVVVAIPALAREFVSGGLASIVTHQV
TNLPGIVLLGLTGALSLP"
complement(4480..4483)
/note="possible RBS upstream of Rv0658c"
complement(4753..5061)
/gene="Rv0659c"
complement(4753..5061)
/gene="Rv0659c"
complement(4753..5061)
/note="Rv0659c" (MTCI376.17), len: 102; unknown, similar
to YW28_MYCTU_Q10867 hypothetical 12.3 kd protein cy39.28
(114 aa), fasta scores: opt: 144 z-score: 213.2 E():
7.3e-05, 30.8% identity in 107 aa overlap. Also similar to
MTCY09f9.22 (32.7% identity in 101 aa overlap)"
/codon_start=1
/transl_table=11
/product="hypothetical protein Rv0659c"
/protein_id="CAB09387.1"
/db_xref="GI:2143302"
/db_xref="SPTREMBL:O06780"
/translation="MRGELWFAATPGGDRPVLVLRDPVADRIGAVVVVALTRTRRG
LVSELELAVENRVPSDCVNFNIIHTLPRTAFRRIRITLSPARLHEACQTLRASTGC
complement(5048..5293)
/gene="Rv0660c"
complement(5048..5293)
/gene="Rv0660c"
/note="Rv0660c" (MTCI376.16), len: 81, some similarity to
IAF016485_130 Halobacterium sp: NRC-1 plasm (100 aa),
32.4%identity in 74 aa overlap"
/codon_start=1

Query Match 80.0%; Score 166.4; DB 1; Length 19770;
Best Local Similarity 87.5%; Pred. No. 7.7e-17;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGCGGCTACAGGTCACAGAGAGCTCG 60
Db 10758 TCAAGGAGAGCGCTACGACCTGGCCGGGTAGCGGCTACAGGTCACAGAGAGCTCG 10817
Qy 61 GCCTGACGTCGCGGATCGGATCCAGCTGCACGCTGACCGGAGGAGACGTCGCGCCA 120
Db 10818 GGCTGATGTCGCGAGGCCATCAGCTGCTGCACGCTGACCGGAGGAGACGTCGCGCCA 10877
Qy 121 CCATCAGTACCTGTCGCGCTGACGAGGCGCCAGCACAGATGACCGTCGCGGCGCGCA 180
Db 10878 CCATCAATATCTGTCGCTGTCAGAGGTCAGACCCAGATGACCGTTCGCGGCGCG 10937
Qy 181 CCGAGGTGCGGTTGAGACCGACGACAT 208
Db 10938 TCAGGTGCGGTTGAGAACCCGACGACAT 10965

RESULT 15
MSU24494
LOCUS MSU24494 3752 bp DNA linear BCT 02-MAR-2000
DEFINITION Mycobacterium smegmatis DNA polymerase (rpoB) gene, complete cds.
ACCESSION U24494
VERSION U24494.1 GI:790347
KEYWORDS

Myobacterium smegmatis.
Myobacterium smegmatis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium.
1 (bases 1 to 3752)
Hetherington,S.V., Watson,A.S. and Patrick,C.C.
Sequence and analysis of the rpoB gene of Mycobacterium smegmatis
Antimicrob. Agents Chemother. 39 (9), 2164-2166 (1995)
96050766
MEDLINE
PUBMED 8540740
2 (bases 1 to 3752)
Hetherington,S.V.
Direct Submission
Submitted (11-APR-1995) Seth V. Hetherington, Infectious Diseases,
St. Jude Children's Research Hospital, 332 N. Lauderdale, Memphis,
TN 38101, USA
Location/Qualifiers
l..3752
/organism="Mycobacterium smegmatis"
/db_xref="taxon:1772"
194..3703
/gene="rpoB"
194..3703
/gene="rpoB"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="DNA polymerase"
/protein_id="AA91426.2"
/db_xref="GI:7144498"
/translation="MLEGCILAVSSQSKSNATINNSVRGPNGVSFAKUREPLEVRGL
LDVQTSFMLVGRSRRVQAADIRGEENPVGLEEVLAEISPIEDFSGMSLSFSDPR
FDVQASDECKDKMTAAFLPFAETFINNTGETKTSOTVFMGPFMGSTGFTIIN
GTRVVSOLVRSPGVYFDETDKSTKTLHSVKVTPGRGAMLEDDVDKRDVGVRIID
RKRPQVLLIKALGNTQEIIVERGFSIMHGTLEKDTSGTDEALLDIYKLRPGE
PPKESAQTLLENLFKERRKIDLARVGRYKNNKGLNAGKPTTSTLTEDVYATIE
YLVRHGEQTSMTVPGVFEVESHDIHGNRRRLTUGELIONQIRVLSRHRVVR
ERM7TDQVEAITPQTLINIRPVVAAIKKEFFTQSQFNDQNNPLSLGTHKRLSALG
PGLSRERAGLEVDRVHPSHYRMCPETPEGNIGLISLSVYARVNPFGFTETPYR
KVVEGVYTDQIDLTADRDHVAQASPTDENGFTEDRVYMRKKGVEFVSADQ
VDYMDVSPQMYSVATAMIPFLEHDDANRALGANNQROAVPLVSEAPLVCTGHLEL
AATDAVTSHTKGTVEEVSADYITVHADDGTQSYRLRKKFARSHNGTCANRPVDA
QQRVEAGVIADGCTQNGEMALKNLLVAIMPVGHNTYEDAIILSNRLVEEDVLTSI
HIEEHIDARDTKLGAEEITRVIPKLSDEVLADLPRIVRIGAEVREDEILVGVTP
KGTELTPEERLIRATFGKAREVRDTSKLVPHGESKGVIGIRVFSREDDDELPGVGN
ELVRYVQAQRKISDGDKLARRHNGVTKLPVEDMPFLPDGTPVDIILNTHGCRV
VLNIGOILETHLGAKAGWNIDVLAGVPDMASKLPEELYSAPADSTVATPVQGAQE
GELAGLIGSTLPNRDGEVWVNDGKATLFDGRSSGEPFPYVTVGVYMLKLHLVDD
KHARSTGPTSMITQPLGKAQFGGREGMECWAMQAGYATILQELLTINSDDTV
GRVYVEAIVKGENIPEGIPESFVKYLLKELQSLCNLNEVLSDSGRAIEMRDGDDDL
ERAAANLGINLSRESASVEDLA"
BASE COUNT 706 a 1225 c 1210 g 611 t
ORIGIN
Query Match 76.9%; Score 160; DB 1; Length 3752;
Best Local Similarity 85.6%; Pred. No. 1e-15;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 1 TCAAGGAGAGCGCTACGACCTGCCCGGTAGCGGCTACAGGTCACAGGTCACAGAGCTCG 60
Db 1068 TCAAGGAGAGCGCTACGACCTGCCCGGTAGCGGCTACAGGTCACAGAGCTCG 1127
Qy 61 GCCTGACGTCGCGGATCCGATCACCAGCTCCAGCTGACCGAGGAGAGAGCTCGTCGCCCA 120
Db 1128 GCCTGAACGCGGCAACCGCATCACCAGCTCGACCTGACCGAGAGAGCTCGTCGCCGA 1187
Qy 121 CCATCAGTACCTGTCGCGCTCCGCTCCAGCGCCAGCAGACATGACCGTCCCGGGCGGCA 180
Db 1188 CCATCAGTACCTGTCGCTCCGCTCCAGCGGCTCAGACCTCGACCGTCCCGGGTGGCG 1247
Qy 181 CCGAGGTGCGGTTGAGACCGACGACAT 208

Db 1248 TCGAGTTCGCGTCGAGTCGACGACAT 1275

Search completed: November 13, 2002, 01:25:04
Job time : 675.723 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 12, 2002, 15:56:36 : Search time 27.8936 Seconds
(without alignments)
2286.857 Million cell updates/sec

Title: US-09-697-123b-8
Perfect score: 208
Sequence: 1 tcaagagaagcgctacgac.....ccggtggaaccgacgacat 208

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	166.4	80.0	970	1	US-08-250-030-1
2	166.4	80.0	970	5	PCT-US95-06790-1
3	142.4	68.5	3447	2	US-08-313-185-57
4	142.4	68.5	3447	3	US-09-082-614A-57
5	39.4	18.9	1894	4	US-09-329-350-32
6	39.4	18.8	1100	3	US-09-248-335-53
7	38	18.3	2742	4	US-09-232-468A-1
8	38	18.3	4897	6	5196516-7
9	36.8	17.7	4411529	4	US-09-103-840A-1
10	36.2	17.4	2634	1	US-08-196-218-31
11	36.2	17.4	2634	1	US-08-681-953-31
12	35.8	17.2	734	4	US-09-221-017B-1070
13	34.8	16.7	4403765	4	US-09-103-840A-2
14	34.6	16.6	1366	6	US-09-280-116-120
15	34.2	16.4	20235	1	US-07-642-734C-3
16	34.2	16.4	20235	3	US-08-439-009A-3
17	34	16.3	1208	2	US-08-403-852D-4
18	34	16.3	1208	3	US-08-510-648B-4
19	34	16.3	1208	4	US-09-231-818-4
20	34	16.3	5392	2	US-08-403-852D-1
21	34	16.3	5392	3	US-08-510-646B-1
22	34	16.3	5392	4	US-09-231-818-1
23	33.8	16.2	6085	4	US-09-029-603-4
24	33.6	16.2	1272	3	US-09-036-987A-28
25	33.6	16.2	1272	4	US-09-370-700-28
26	33.4	16.1	1112	4	US-09-434-288-9
27	33.4	16.1	1281	4	US-09-105-537-19

28	33.4	16.1	13613	4	US-09-105-537-3	Sequence 3, Appl
c 29	33.4	16.1	30001	1	US-08-125-468-1	Sequence 1, Appl
c 30	33.4	16.1	30001	2	US-08-474-933-1	Sequence 1, Appl
31	33.4	16.1	38506	3	US-09-320-878-19	Sequence 19, Appl
32	33.4	16.1	44377	2	US-08-804-227C-7	Sequence 7, Appl
33	33.4	16.1	44377	2	US-08-804-198-1	Sequence 1, Appl
c 34	33.2	16.0	49272	1	US-08-614-770A-1	Sequence 1, Appl
35	33	15.9	765	3	US-08-718-904-79	Sequence 79, Appl
36	33	15.9	1722	4	US-09-385-028-15	Sequence 15, Appl
37	33	15.9	11604	4	US-09-385-028-13	Sequence 13, Appl
38	33	15.9	15079	4	US-09-385-028-1	Sequence 1, Appl
39	32.6	15.7	3048	1	US-08-188-228-47	Sequence 47, Appl
40	32.6	15.7	3048	1	US-08-332-643-41	Sequence 41, Appl
41	32.6	15.7	3048	1	US-08-332-638-47	Sequence 47, Appl
42	32.4	15.6	3435	1	US-08-366-577-1	Sequence 1, Appl
43	32.4	15.6	3435	5	PCT-US96-00005-1	Sequence 1, Appl
44	32.4	15.6	11219	1	US-07-642-734C-1	Sequence 1, Appl
45	32.4	15.6	11219	3	US-08-439-009A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-250-030-1
; Sequence 1, Application US/08250030
; Patent No. 5643723
; GENERAL INFORMATION:
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: Detection of a Genetic Locus Encoding
; TITLE OF INVENTION: Resistance to Rifampin in Mycobacterial Cultures and in
; TITLE OF INVENTION: Clinical Specimens
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg & Woessner
; STREET: 3500 IDS Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/250,030
; FILING DATE: 26-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Muetting, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 150.105US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-250-030-1

Query Match 80.0%; Score 166.4; DB 1; Length 970;
Best Local Similarity 87.5%; Pred. No. 3.7e-33;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 1 TCAAGGAGAGCGCTACGACCTGGCCCGGTGTGGCGGATACAGGTCAACAAGAGCTGG 60
|||||
Db 26 TCAAGGAGAGCGCTACGACCTGGCCCGGTGTGGCGGATACAGGTCAACAAGAGCTGG 85
|||||

QY 61 GCGTGAACCAATCATCGGATCAGCACGACGCTGACCGAAGAAGACGCTCGCGCA 120
| | | | |
Db 86 GCGTGCATGCGGCGGCCATCACGTCGTCACGCTGACCGAAGAAGACGCTCGTGCCA 145
QY 121 CCATCGAGTATCTGGTCCGCTTGCAGGAGGCCAGCCACGATGACCGTGCCTGGCGGGG 180
| | | | |
Db 146 CCATCGAATATCTGGTCCGCTTGCAGGAGGTCAGACACGATGACCGTTCGCGGGCGG 205
QY 181 TCAGGTCGCGCTGGAAACCGACGACAT 208
| | | | |
Db 206 TCAGGTCGCGCTGGAAACCGACGACAT 233

RESULT 2
PCT-US95-06790-1
: Sequence 1, Application PC/TUS9506790
: GENERAL INFORMATION:
: APPLICANT: Mayo Foundation for Medical Education and Research
: APPLICANT: and Hoffmann-La Roche Inc.
: TITLE OF INVENTION: Detection of a Genetic Locus Encoding
: TITLE OF INVENTION: Resistance to Rifampin
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Schwegman, Lundberg & Woessner
: STREET: 3500 IDS Center
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/06790
: FILING DATE: 26-MAY-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Raesch, Kevin W.
: REGISTRATION NUMBER: 35,651
: REFERENCE/DOCKET NUMBER: 150.105W01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-339-0331
: TELEFAX: 612-339-3061
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 970 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
PCT-US95-06790-1

Query Match 80.0% Score 166.4; DB 5; Length 970;
Best Local Similarity 87.5% Pred. No. 3.7e-33;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTGCGCGCTGCGCGCGATACACAGGTCAACAAAGCTGG 60
| | | | |
Db 26 TCAAGGAGAGCGCTACGACCTGCGCGCGCTGCGCGCTATATAGGTCAACAAAGCTGG 85
QY 61 GCGTGAACCAATCATCGGATCAGCACGACGCTGACCGAAGAAGACGCTCGTGCCA 120
| | | | |
Db 86 GCGTGCATGTCGGGAGCCCATCACGTCGTCGACGCTGACCGAAGAAGACGCTCGTGCCA 145
QY 121 CCATCGAGTATCTGGTCCGCTTGCAGGAGGCCAGCCACGATGACCGTGCCTGGCGGGG 180
| | | | |
Db 146 CCATCGAATATCTGGTCCGCTTGCAGGAGGTCAGACACGATGACCGTTCGCGGGCGG 205
QY 181 TCAGGTCGCGCTGGAAACCGACGACAT 208
| | | | |
Db 206 TCAGGTCGCGCTGGAAACCGACGACAT 233

RESULT 3
US-08-313-185-57
: Sequence 57, Application US/08313185
: Patent No. 5851763
: GENERAL INFORMATION:
: APPLICANT: Heym, Beate
: APPLICANT: Cole, Stewart
: APPLICANT: Young, Douglas
: APPLICANT: Zhang, Ying
: APPLICANT: Honore, Nadine
: APPLICANT: Telenti, Amelio
: APPLICANT: Bodmer, Thomas
: TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
: TITLE OF INVENTION: in Mycobacterium Tuberculosis
: NUMBER OF SEQUENCES: 66
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSEE: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/313,185
: FILING DATE: 12-OCT-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Meyers, Kenneth J.
: REGISTRATION NUMBER: 25,146
: REFERENCE/DOCKET NUMBER: 02356.0068-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 408-4000
: TELEFAX: (202) 408-4400
: INFORMATION FOR SEQ ID NO: 57:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3447 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
US-08-313-185-57

Query Match 68.5% Score 142.4; DB 2; Length 3447;
Best Local Similarity 80.3% Pred. No. 4.1e-27;
Matches 167; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTGCGCGCTGCGCGCGATACAAAGTCAACAAAGCTGG 60
| | | | |
Db 809 TCAAGGAGAGCGCTACGACCTGCGCGCGCTGCGCGCGATACAAAGTCAACAAAGCTGG 868
QY 61 GCGTGAACCAATCATCGGATCAGCACGACGCTGACCGAAGAAGACGCTCGTGCCA 120
| | | | |
Db 869 GGTGTCACGCGCGGTGAGTGTGATCAGCTGCTGCACCGTGCACCGAAGAGGATGTCGTCGCCA 928
QY 121 CCATCGAGTATCTGGTCCGCTTGCAGGAGGCCAGCCACGATGACCGTGCCTGGCGGGG 180
| | | | |
Db 929 CCATGAGTATCTGGTTCGCTGTCATGAGGTCAGTCGACATGACTGTCCCGAGGTGGG 988
QY 181 TCAGGTCGCGCTGGAAACCGACGACAT 208
| | | | |
Db 989 TAGAAGTCCAGTGGAAACCTGACGATAT 1016

RESULT 4
US-09-082-614A-57

Sequence 57, Application US/09082614A
Patent No. 6124098
GENERAL INFORMATION:
APPLICANT: Heym, Beate
APPLICANT: Cole, Stewart
APPLICANT: Young, Douglas
APPLICANT: Zhang, Ying
APPLICANT: Honore, Nadine
APPLICANT: Telenti, Amalio
APPLICANT: Rodmer, Thomas
TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESS: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,614A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/313,185
FILING DATE: 12-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 02356.0068-00000
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 3447 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-082-614A-57

Query Match 68.5%; Score 142.4; DB 3; Length 3447;
Best Local Similarity 80.3%; Pred. No. 4.1e-27;
Matches 167; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 1 TCAGGAGGACGGCTACGACCTGGCCGCTCGGCCGATACAAAGTCAACGAAGCTGG 60
Db 809 TCAGGAGGACGGCTACGACCTGGCCGCTCGGCCGATACAAAGTCAACGAAGCTGG 868
QY 61 GCCTGAACACCAATCATCCGATCCACACGACGCTGACCGAAGAGACGCTCGCCCA 120
Db 869 GGTTCACCGCGGTGAGTGTATCAGCTGCTCCAGCTGACCGAAGAGGATGCTCGCCCA 928
QY 121 CCATCGAGTATCTGGTCCGCTGACGACGCGCCAGGCCACGATGACGCTCCGGGGGGG 180
Db 929 CCATAGAGTACCTGGTTCGCTCATGAGGCTGAGTCAACATGACTGTCCAGGTGGGG 988
QY 181 TCAGAGTGCCTGGTGAACACGACGACAT 208
Db 989 TAGAAGTCCAGTGGAAACTGACGATAT 1016

RESULT 5
US-09-329-350-32
Sequence 32, Application US/09329350

Patent No. 6184019
GENERAL INFORMATION:
APPLICANT: Miettinen-Oinonen, Arja
APPLICANT: Londesborough, John
APPLICANT: Vehmaanper, Jari
APPLICANT: Haakana, Heli
APPLICANT: M ntyl , Arja
APPLICANT: Lantto, Raija
APPLICANT: Elovainio, Minna
APPLICANT: Joutsjoki, Vesa
APPLICANT: Paloheimo, Marja
APPLICANT: Suominen, Pirkko
TITLE OF INVENTION: NOVEL CELLULASES, THE GENES ENCODING THEM AND
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/329,350
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/841,636
FILING DATE: 30-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/005,335
FILING DATE: 17-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/007,926
FILING DATE: 04-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,840
FILING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/732,181
FILING DATE: 16-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/PI96/00550
FILING DATE: 17-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Shea Jr., Timothy
REGISTRATION NUMBER: 41,306
REFERENCE/DOCKET NUMBER: 1716.0510006/MAC/TJS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 1894 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Melanocarpus albomyces
STRAIN: ALKO4237
FEATURE:
NAME/KEY: exon
LOCATION: 233..838
OTHER INFORMATION: /product= "50K-cellulase"
FEATURE:
NAME/KEY: exon
LOCATION: 916..1596


```
; OTHER INFORMATION: /product- "50K-cellulase"
US-09-329-350-32
Query Match      18.9%; Score 39.4; DB 4; Length 1894;
Best Local Similarity 55.5%; Pred. No. 0.15;
Matches 76; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

U; 56 GCTGGGCTGAACACCAATCATCCGATACACGACGACGCTGACCGAAGAAGACGCTCGT 115
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1131 GTTCCGCTGCACACGCGGCGCTTACCGTGTGTGACGCACTTCCCGCGCAGCCGCA 1190

QY 116 CGCCACCATCGAGTATCTGCTCGCCTGACAGAGGCCAGGCCACCATGACCGTGCCTGG 175
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1191 GGGCCGGCTCGAGAGCATCACCGCTGTACGTGCGAGGACGGCAAGTGATCGAGTCGTA 1250
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 176 CGGGGTGAGGTGCCGG 192
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1251 CGTCGTGAGCGCCCG 1267

RESULT 6
US-09-248-335-53
; Sequence 53, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEF, DANIEL
; TITLE OF INVENTION: PLANT GLUTAMINOLONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1128-A
; CURRENT APPLICATION NUMBER: US/09/248,335
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 08/924,759
; EARLIER FILING DATE: 1997-September-05
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 53
; LENGTH: 1100
; TYPE: DNA
; ORGANISM: maize
US-09-248-335-53

Query Match      18.8%; Score 39; DB 3; Length 1100;
Best Local Similarity 52.8%; Pred. No. 0.18;
Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 47 CAACAAGAAGCTGGGCTGAACACCAATATCTGCTCCGCTGACACGACGACGCTGACCGAAGA 106
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 193 CAACAGCGCGCTCTCTGGGATCCAAACCGGTGCAACAGAGCGTCCGCTGCTCTCTCCA 252

QY 107 AGAGTCGTGCGCCACCATCATCTGCTGCTCCGCTGACGAGGCGCCAGCCAGCATGAC 166
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 253 CGCGCGCGCGCCATAACAGAGTCCCAGGTATCTCTGACATCATCGACGAGGTGTGGGC 312

QY 167 CGTCCCGCGGCGGCTCGAGCTGCGGTGGGAACCGAGCA 205
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 313 GGGGACGCGGCGCGCTGCTCCGCGCGACCCCTAGCA 351

RESULT 7
US-09-232-468A-1
; Sequence 1, Application US/09232468A
; Patent No. 6207165
; GENERAL INFORMATION:
; APPLICANT: AUDONNET et al.
; TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE FORMULA AGAINST PORCINE
; TITLE OF INVENTION: REPRODUCTIVE AND RESPIRATORY PATHOLOGIES
; FILE REFERENCE: 454313-2230
; CURRENT APPLICATION NUMBER: US/09/232,468A
; CURRENT FILING DATE: 1999-01-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2742

; TYPE: DNA
; ORGANISM: Pseudorabies virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2742)
US-09-232-468A-1

Query Match      18.3%; Score 38; DB 4; Length 2742;
Best Local Similarity 60.8%; Pred. No. 0.36;
Matches 62; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 62 CCTGAACACCAATCATCCGATCACCACGACGCTGACCCGAAGAAGACGCTCGTCGCCAC 121
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2484 CATGAAGGCCCTGTACCCCGCTCACCAGCAAGACGCTCAAGGAGGACGCGCTCGACCAAGG 2543

QY 122 CATCGAGTATCTGGTCCGCTGCAGGAGGCCAGGCCACGAT 163
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2544 CGACGTGGAGGAGGCCAAGCTGGACAGCCGCCGGACATGAT 2585

RESULT 8
5196516-7
; Patent No. 5196516
; APPLICANT: SCHREURS, CHRISTA S.;METTENLEITER, THOMAS C.
; SIMON, ARTUR J.;LUKAS, NOEMI;RZIHA, HANNS J.
; TITLE OF INVENTION: PSEUDORABIES VIRUS VACCINE
; NUMBER OF SEQUENCES: 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/383,833
; FILING DATE: 21-JUL-1989
; SEQ ID NO:7;
; LENGTH: 4897
5196516-7

Query Match      18.3%; Score 38; DB 6; Length 4897;
Best Local Similarity 60.8%; Pred. No. 0.39;
Matches 62; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 62 CCTGAACACCAATCATCCGATCACCACGACGCTGACCCGAAGAAGACGCTCGTCGCCAC 121
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3355 CATGAAGGCCCTGTACCCCGCTCACCAGCAAGACGCTCAAGGAGGACGCGCTCGACCAAGG 3414

QY 122 CATCGAGTATCTGGTCCGCTGCAGGAGGCCAGGCCACGAT 163
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3415 CGACGTGGAGGAGGCCAAGCTGGACAGCCGCCGGACATGAT 3456

RESULT 9
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match      17.7%; Score 36.8; DB 4; Length 4411529;
Best Local Similarity 55.5%; Pred. No. 1.9;
Matches 71; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
```

```
QY 50 CAGAGAGCTGGCCCTGACACCAATCATCCGATCACCACGACGAGCTGACCGGAAGA 109
|| || || || || || || || || || || || || || || || || || || || || ||
Db 595927 CACCAGATGACCTGGACCGCGGCCAGCAGGTATGGCTGATCACCCTACCCCATACGA 595868
|| || || || || || || || || || || || || || || || || || || || || ||
QY 110 GCTGCTGCCACCACTCAGTATCTGCTCGGCTGCACGAGAGGCCAGGCCAGATGACCGT 169
|| || || || || || || || || || || || || || || || || || || || || ||
Db 595867 ACTGCGGCCACCATCGCCCGCGGCTCGGCTAATCGGCTGGGACCGCTGCGCGGA 595808
|| || || || || || || || || || || || || || || || || || || || || ||
QY 170 GCGGCGCG 177
|| || || || || || || || || || || || || || || || || || || || || ||
Db 595807 GTCGCTCG 595800
|| || || || || || || || || || || || || || || || || || || || || ||

RESULT 10
US-08-196-218-31
: Sequence 31, Application US/08196218
: Patent No. 5614619
: GENERAL INFORMATION:
: APPLICANT: Piepersberg, Wolfgang
: APPLICANT: Stockmann, Michael
: APPLICANT: Taleghani, Kampiz Mansouri
: APPLICANT: Distler, Jurgen
: APPLICANT: Grabley, Susanne
: APPLICANT: Sichel, Petra
: APPLICANT: Brau, Barbara
: TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes
: TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSEE: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: United States
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/196,218
: FILING DATE: 25-AUG-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Ogden, Stasia L.
: REGISTRATION NUMBER: 36,228
: REFERENCE/DOCKET NUMBER: 02481.1372-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-408-4000
: TELEFAX: 202-408-4400
: INFORMATION FOR SEQ ID NO: 31:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2634 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 3..401
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 416..1531
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1561..2625
US-08-196-218-31
```

Query Match 17.4% Score 36.2: DB 1: Length 2634;

```
Best Local Similarity 49.2%: Pred. No. 1:
Matches 95: Conservative 0: Mismatches 98: Indels 0: Gaps 0:
QY 12 CGGTACGACCTGGCCCGCTGTCGGCCGATACAAGGTCAACAAGAAGCTGGCGCTGAACACC 71
|| || || || || || || || || || || || || || || || || || || || || ||
Db 2197 CTCTACGACCGCTACTGGAAGGACACCGGGAGGGTGCAGGACGCTCTTGAGTCGACAGC 2256
|| || || || || || || || || || || || || || || || || || || || || ||
QY 72 AATCATCCGATCACCAGCAGCAGCGTGCAGCAAGAAGACGCTGTCGCCACCATCGAGTAT 131
|| || || || || || || || || || || || || || || || || || || || || ||
Db 2257 CAGCTCTGACGCGCTGACCGCGCGCTGCAGCGACAGGTGCAGCGCCGACAGCGTCTC 2316
|| || || || || || || || || || || || || || || || || || || || || ||
QY 132 CTGCTCGCGCTGCAGGAGGCCAGGCTACCATGATGACTGTCGGCGCGGGTGCAGGTCGCG 191
|| || || || || || || || || || || || || || || || || || || || || ||
Db 2317 GTCGGCGGGTGTGATCGAGCGCGGGCGCGCATCGTCGCGGGTCGAGGGCGCG 2376
|| || || || || || || || || || || || || || || || || || || || || ||
QY 192 GTGGAACCGACG 204
|| || || || || || || || || || || || || || || || || || || || || ||
Db 2377 GCGATCATCGCG 2389
|| || || || || || || || || || || || || || || || || || || || || ||

RESULT 11
US-08-681-953-31
: Sequence 31, Application US/08681953
: Patent No. 5710032
: GENERAL INFORMATION:
: APPLICANT: Piepersberg, Wolfgang
: APPLICANT: Stockmann, Michael
: APPLICANT: Taleghani, Kampiz Mansouri
: APPLICANT: Distler, Jurgen
: APPLICANT: Grabley, Susanne
: APPLICANT: Sichel, Petra
: APPLICANT: Brau, Barbara
: TITLE OF INVENTION: Secondary-Metabolite Biosynthesis Genes
: TITLE OF INVENTION: From Actinomycetes, Method of Isolating Them, and Their
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSEE: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: United States
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/681,953
: FILING DATE: 30-JUL-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/196,218
: FILING DATE: 25-AUG-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Ogden, Stasia L.
: REGISTRATION NUMBER: 36,228
: REFERENCE/DOCKET NUMBER: 02481.1372-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-408-4000
: TELEFAX: 202-408-4400
: INFORMATION FOR SEQ ID NO: 31:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2634 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 3..401
```


Db 597168 GTCGCTCG 597161

RESULT 14

US-09-280-116-120
Sequence 120, Application US/09280116A
Patent No. 6331427

GENERAL INFORMATION:

APPLICANT: Robison, Keith E.
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
FILE REFERENCE: 5800-24, 035900/176965
CURRENT APPLICATION NUMBER: US/09/280,116A
CURRENT FILING DATE: 1999-03-26
NUMBER OF SEQ ID NOS: 268
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 120
LENGTH: 1366

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: zinc proteases

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)..(1366)

OTHER INFORMATION: n = a, t, c or g

US-09-280-116-120

Query Match

Rest Local Similarity 16.6%; Score 34.6; DB 4; Length 1366;

Matches 67; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 59 GGGCCGTGAACACCAATCATCCGATCACCAGCAGCGTGTGACCGAAGAGCGTGTGCG 118

||||| ||||| || || || || || || || || || || || || || || || || || || ||

Db 309 GGGCCTGGCCACCTATGGACCCGCTCGCCACCGAGACCTACGGTGTGCTTTCAC 368

||||| ||||| || || || || || || || || || || || || || || || || || || ||

QY 119 CACCATCGAGTATCTGCTCCGCTGCACGAGCGGCCAGCGATCACCGTCCCGGCGG 178

|| || || || || || || || || || || || || || || || || || || || || || || || ||

Db 369 CTGCCCTGGAGACTGCTTCGGCTGGAGCGCCCTGCACCGCGGATGAAGCTNCTGGGAGA 428

|| || || || || || || || || || || || || || || || || || || || || || || || ||

QY 179 GGTTC 182

|| ||

Db 429 GCAC 432

|| ||

RESULT 15

US-07-642-734C-3
Sequence 3, Application US/07642734C
Patent No. 5824513

GENERAL INFORMATION:

APPLICANT: Katz, L
APPLICANT: Donadio, S
TITLE OF INVENTION: Recombinant DNA Method for Producing
TITLE OF INVENTION: Erythromycin Analogs
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:

ADDRESSEE: Edward H. Gorman

STREET: Abbott Laboratories D377/AP6D-2 One Abbott

CITY: Park Rd

STATE: IL

COUNTRY: US

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/642,734C

FILING DATE: 17-JAN-91

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Danckers, Andreas M
REGISTRATION NUMBER: 32652
REFERENCE/DOCKET NUMBER: 4952-US-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9396
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 20235 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Saccharopolyspora erythraea
STRAIN: NRRL 238
FEATURE:
NAME/KEY: CDS
LOCATION: 19..10722
OTHER INFORMATION: /codon_start= 19
OTHER INFORMATION: /function= "gene eryA"
OTHER INFORMATION: /product= "eryA ORF2 encoding modules 3 & 4 for
OTHER INFORMATION: 6-deoxyerythronolide B"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19..4470
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 97..1482
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacyl ACP synthase of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1693..2670
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3406..3921
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4171..4428
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain of module 3"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4471..10722
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 4"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4471..5847
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacylACPsynthase domain of module"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6054..7026
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain of module 4"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7165..9216
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: dehydratase and enoylreductase domains m"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9433..9984

OTHER INFORMATION: /function= "approximate span
FEATURE: beta-ketoreductase of module 4"
NAME/KEY: misc_feature
LOCATION: 10225..10483
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain of module 4"
FEATURE:
NAME/KEY: CDS
LOCATION: 10723..20235
OTHER INFORMATION: /codon_start= 10723
OTHER INFORMATION: /function= "gene-eryA"
OTHER INFORMATION: /product= "orf3 encoding modules 5 & 6
OTHER INFORMATION: 6-deoxyerythronolide B formatio"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 10723..15165
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 5"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 10831..12174
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12379..13350
OTHER INFORMATION: /function= "approximatr span of
OTHER INFORMATION: acyltransferase domain of module 5"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 14062..14610
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase of module 5"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 14857..15114
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain of module 5"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15166..20235
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: module 6"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 15172..16569
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoacylACPSynthase domain of modul"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 16768..17721
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyltransferase domain of module 6"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 18379..18921
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: beta-ketoreductase domain of module 6"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19149..19398
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: acyl carrier domain of module 6"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 19492..20235
OTHER INFORMATION: /function= "approximate span of
OTHER INFORMATION: thioesterase domain of module 6"
US-07-642-734C-3

Query Match 16.4%; Score 34.2; DB 1; Length 20235;
Best Local Similarity 52.4%; Pred. No. 4.2;

Matches 75; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
Oy 64 TGAACACCAATCATCCGATCACACGACGCGTGGACCGAAGAAGACGTCTCGTCCACCA 123
Db 17573 TGAGCCCGCACCCGGTGTCTACCGCGGGGGTGCAGGAGATCGCCGGGACGCCGTGGCCA 17632
Oy 124 TCGAGTATCTGGTCCGGCTGCACGAGGGCCAGGCCACGATGACCGTGCCTGGGGGGGGTFCG 183
Db 17633 TCGGTGCTGCACCGCGACACCGCGGAGGAGCACCTGATCGCGGAGCTCGCCGGGCGC 17692
Oy 184 AGGTGCCGGTGGAAACCGACGAC 206
Db 17693 ACGTGCACGGGTGGCCCGTGGAC 17715

Search completed: November 12, 2002, 20:14:08
Job time : 5353.89 secs

Db	1638	GCGCACGATGGTTT---	GATGACTCTTACTGAAGAGACATCGCAACCACCATCGAG	1694
Oy	129	TATTCTGGTCGCGCTGCACGAGGCCAGCCAGTACCGCTCCGGGGCGGGTGCGAGGTG	188	
Db	1695	TACCTGGTGGCTGCAACGAGGTGACGCGGTCTATGACTTCTCCAATGGTGAAGAGATC	1754	
Oy	189	CCGGTGGAAACGACGACAT	208	
Db	1755	CCAGTCGACAGCCGATGACAT	1774	

```

RESULT 5
US-09-887-052-5
: Sequence 5, Application US/09887052
: Patent No. US20020119537A1
: GENERAL INFORMATION:
: APPLICANT: MOECKEL, Bettina
: APPLICANT: BATHE., Brigitte
: APPLICANT: HERMANN, Thomas
: APPLICANT: PFEFFERLE, Walter
: APPLICANT: BINDER, Michael
: TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE
: FILE REFERENCE: 204212US0X
: CURRENT APPLICATION NUMBER: US/09/887,052
: CURRENT FILING DATE: 2001-06-25
: PRIOR APPLICATION NUMBER: DE10107229.5
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 5
: LENGTH: 5099
: TYPE: UNA
: ORGANISM: Corynebacterium glutamicum
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (702)..(4196)
US-09-887-052-5

```

```

RESULT 6
US -10-062-254-109
: Sequence 109, Application US/10062254
: Patent No. US2002013882A1
: GENERAL INFORMATION:
: APPLICANT: Cahoone, Edgar B
: APPLICANT: Cahoone, Rebecca E
: APPLICANT: Falco, Saverio Carl
: APPLICANT: Fang, Yiwen
: APPLICANT: Hantke, Sabine S.
: APPLICANT: Lee, Jian-Ming
: APPLICANT: Li, Zhongsen
: APPLICANT: Miao, Guo-Hua

```

```

: APPLICANT: Morgante, Michele
: APPLICANT: Niu, Xiping
: APPLICANT: Odell, Joan
: APPLICANT: Rafalski, Antoni
: APPLICANT: Sakai, Hajime
: APPLICANT: Zheng, Peizhong
: APPLICANT: Zhu, Qun
: TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved in Plant Metabolism
: FILE REFERENCE:
: CURRENT APPLICATION NUMBER: US/10/062,254
: CURRENT FILING DATE: 2002-02-01
: PRIOR APPLICATION NUMBER: 09/630,346
: PRIOR FILING DATE: 2000-07-28
: PRIOR APPLICATION NUMBER: 60/146511
: PRIOR FILING DATE: 1999-07-30
: PRIOR APPLICATION NUMBER: 60/156006
: PRIOR FILING DATE: 1999-09-23
: PRIOR APPLICATION NUMBER: 60/156899
: PRIOR FILING DATE: 1999-09-30
: PRIOR APPLICATION NUMBER: 60/157287
: PRIOR FILING DATE: 1999-10-01
: PRIOR APPLICATION NUMBER: 60/169767
: PRIOR FILING DATE: 1999-12-09
: PRIOR APPLICATION NUMBER: 60/171054
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: 60/172958
: PRIOR FILING DATE: 1999-12-21
: PRIOR APPLICATION NUMBER: 60/171515
: PRIOR FILING DATE: 1999-12-22
: PRIOR APPLICATION NUMBER: 60/173535
: PRIOR FILING DATE: 1999-12-29
: NUMBER OF SEQ ID NOS: 375
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 109
: LENGTH: 704
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (640)
: NAME/KEY: unsure
: LOCATION: (663)
: NAME/KEY: unsure
: LOCATION: (688)
: NAME/KEY: unsure
: LOCATION: (692)
: NAME/KEY: unsure
: LOCATION: (696)
: US-10-062-254-109

Query Match 18.1%; Score 37.6; DB 12; Length 704;
Best Local Similarity 56.5%; Pred. No. 0.14;
Matches 70; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 4 AGGAGAGCGCTACGACCTGGCCCGTCTCGGCCGATACAAAGGTCAACAAGCTGGCC 63
Db 275 AGCTGATGGCCGACCTGGCCACCCATCGCGCAAGGAGGTCAACCAAGGTGGCC 334
Qy 64 TGAACACCAATCATCCGATCACCACGACGCGTGCACGAAGAAGACGTCGTCGCCACCA 123
Db 335 TCGTCTACCAGCTCAACATCGCCCCCAAGAAGATCGCGCGTTGATGAGGAGATCTTCGTGG 394
Qy 124 TCGA 127
Db 395 TCGA 398

RESULT 7
US-10-062-254-111
: Sequence 111, Application US/10062254
: Patent No. US2002013882A1
: GENERAL INFORMATION:
: APPLICANT: Cahoon, Edgar B

```


Qy	17	CGACCTGCCCGTGTGCGCCGATCAAGGTCACAGAAGCTGGCGCTCAACACCAATCA	76
Db	15	CGCCAGTGGCCGCCACCGCAGCGAGGACGACACTCGCGGAGCCTGCCGGGCACTA	74
Qy	77	TCCGATACCCACGACGACGCTGACCGAAGAGAGCTCTCGCGCCACCATCGAGTATCTCGT	136
Db	75	TCCCTATCACACCGGGATGCGCCGACAGAGTACGAGGAAGGCCAAGACGACCTCGAGAT	134
Qy	137	CGCGCTGCACAGGCGCCAGGCCACCATCACCTGCGCGGCGGGGTCGAGGTG	188
Db	135	CGAAGTCTCAAGGTGACAGCTGGTCAAGSAGACCGGCCAGCCGCTGGTG	186

```

RESULT 9
US-09-815-242-7946
: Sequence 7946, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: TITLE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23

```

```
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7946
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(492)
US-09-815-242-7946
```

```
Query Match          17.8%; Score 37; DB 10; Length 492;
Best Local Similarity 51.5%; Pred. No. 0.19; Mismatches 0; Indels 0; Gaps 0;
Matches 85; Conservative 0;

QY 21 CTGGCCCGTGTGGCCGATACAAAGCTCAACAAGAGCTGGGCTGAACACCAATCATCCG 80
    ||| || ||| || ||| || ||| || ||| || ||| || ||| || ||| || |||
Db 142 CTCACCACCGTGGGCACCATGAGGTGATCGAGCAGATCACCAAGAACCCTCAACAAGCTG 201
    ||| || ||| || ||| || ||| || ||| || ||| || ||| || ||| || |||
QY 81 ATCACCACGACGCGCTGACCGAAGAGAGCTGCTGGCCACCACATCGAGTATCTGGTCCGC 140
    ||| || ||| || ||| || ||| || ||| || ||| || ||| || ||| || |||
Db 202 ATCGAAGTGGTCAAGCTGGTGTCTGTCTCGAAACGCCCATATCGAGCGGAGCTGATG 261
    ||| || ||| || ||| || ||| || ||| || ||| || ||| || ||| || |||
QY 141 CTCACGAGCGGCGAGCGACGATGACGCTGCCGGCGGGCTCGAG 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 CTGTGTAAGGTCAAGCCACGCGGCGCCACGCGCGGAGGTCAAG 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 10
US-09-294-093B-1392
; Sequence 1392, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Laura, Y.
; APPLICANT: Sherman, Bradley, K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; CURRENT FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 1392
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: incyte ID No. US20010051335A1 700344235H1
; NAME/KEY: unsure
; LOCATION: 158
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-1392
```

```
Query Match          17.4%; Score 36.2; DB 10; Length 278;
Best Local Similarity 57.5%; Pred. No. 0.28; Mismatches 48; Indels 0; Gaps 0;
Matches 65; Conservative 0;

QY 4 AGGAGAAGCGCTACGACCTGGCCCGTGTGGCCGATACAAAGGTCAACAAGAAGCTGGGCC 63
    || || ||| || ||| || ||| || ||| || ||| || ||| || ||| || |||
Db 164 AGTGATGCCCGCAGCTGGCCACCGATCGCCGCAAGGAGGTCCACCAAGAAGTGGGCC 223
    || || ||| || ||| || ||| || ||| || ||| || ||| || ||| || |||
QY 64 TGAACACCAATCATCCGATCACCAACGACGCTGACCAAGAGAGCTGCTC 116
    || || || || ||| || ||| || ||| || ||| || ||| || ||| || |||
Db 224 TCGTCTACCAGTCAACATCGCCCCCAAGAAGATCGCGTTGATGAGGATC 276
    || || || || ||| || ||| || ||| || ||| || ||| || ||| || |||
```

RESULT 11

```
US-09-880-107-3678/c
; Sequence 3678, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 3678
; LENGTH: 2693
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X04654
US-09-880-107-3678
```

```
Query Match          17.4%; Score 36.2; DB 10; Length 2693;
Best Local Similarity 52.3%; Pred. No. 0.38;
Matches 80; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 39 TACAAGGTCAACAAGAAGCTGGCGCTGAACACCAATCATCCGATCACCACGACGCTG 98
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 273 TACAAGCTCACCAACAAGCTGGGTACGGCTTCGTATCACCTGATGGCGCGTGGCGGTG 214
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 99 ACCGAAGAAGAGCTGCTGCCACCATCGAGTATCTGGTCCGCTGCACGAGGCCAGGCC 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 213 TGCCTGGTGGCGGTGACCATGCCCTGGTGTGATGCTGCACGCGCTGCTGCGCTCGGCC 154
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 159 ACGATGACGCTGCCGGGGGTGCGAGGTGCCG 191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 153 AACAAAGTTCGTGCCCATCAAGGGCAAGGGCGCG 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 12
US-09-815-242-7738
; Sequence 7738, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

CM nucleic - nucleic search, using sw model
Run on: November 12, 2002, 16:59:42 ; Search time 1109.13 Seconds
(without alignments)
3037.202 Million cell updates/sec

Title: US-09-697-123B-8
Perfect score: 208
Sequence: 1 tcaagagagacgctacac.....ccggtgaaaccagacacat 208

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DR seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_Other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DR ID	Description
c 1	160	76.9	1282	9	AI770311
2	44	21.2	1019	17	CNS0185F
3	43.2	20.8	630	9	AI405818
4	43.2	20.8	662	9	AI405617
5	40.6	19.5	397	12	BF007092
6	40.6	19.5	568	13	BI995346

7	39.2	18.8	516	10	BE056696	BE056696	894011A09
8	39.2	18.8	666	14	BQ817943	BQ817943	1030067C0
9	39.2	18.8	681	14	BQ817941	BQ817941	1030067C0
10	39	18.8	464	13	BM501065	BM501065	PAC000000
11	39	18.8	540	12	BF008145	BF008145	1682830_A
12	39	18.8	710	12	BG321153	BG321153	zm04_0510
13	39	18.8	762	12	BG836339	BG836339	zm06_0180
14	39	18.8	996	11	AY104676	AY104676	zea may
15	39	18.8	1856	11	AY109418	AY109418	zea may
16	38.8	18.7	656	13	BI960536	BI960536	HVSMEn002
17	38.6	18.6	701	12	BG836253	BG836253	zm06_02H0
18	38.4	18.5	413	14	BQ819187	BQ819187	1030076A1
19	38.4	18.5	648	14	BQ579756	BQ579756	WHE2974_B
20	38.4	18.5	877	13	BI957118	BI957118	HVSMEn000
21	38.2	18.4	550	14	BQ767542	BQ767542	EBr008_S0
22	38	18.3	511	10	AV628007	AV628007	AV628007
23	38	18.3	543	10	AV631761	AV631761	AV631761
24	38	18.3	554	10	AV631497	AV631497	AV631497
25	38	18.3	566	9	AL825144	AL825144	AL825144
c 26	38	18.3	725	14	BQ823367	BQ823367	1030110E0
27	37.8	18.2	448	12	BG463919	BG463919	EM1_52_E0
28	37.8	18.2	539	12	BG556957	BG556957	EM1_40_C0
29	37.8	18.2	589	12	BG322881	BG322881	EM1_15_C0
c 30	37.8	18.2	790	17	BH374681	BH374681	AG-ND-170
c 31	37.8	18.2	925	17	CNS0091P	CNS0091P	AL053013
32	37.6	18.1	431	13	BI097733	BI097733	Drosophill
33	37.6	18.1	507	13	BI096671	BI096671	949017G04
c 34	37.6	18.1	531	13	BM348454	BM348454	MEST291-B
c 35	37.6	18.1	579	13	BM340507	BM340507	MEST323-F
c 36	37.6	18.1	586	12	BG842655	BG842655	MEST34-E0
c 37	37.6	18.1	593	13	BM079854	BM079854	MEST100-E
c 38	37.6	18.1	597	13	BM333485	BM333485	MEST156-D
c 39	37.6	18.1	606	13	BM337568	BM337568	MEST209-B
c 40	37.6	18.1	610	13	BM073467	BM073467	MEST67-H1
c 41	37.6	18.1	613	13	BM340956	BM340956	MEST328-C
c 42	37.6	18.1	622	13	BM337575	BM337575	MEST209-C
c 43	37.6	18.1	651	13	BM332944	BM332944	MEST181-G
c 44	37.6	18.1	659	13	BM266851	BM266851	MEST386-H
c 45	37.6	18.1	659	13	BM337566	BM337566	MEST209-H

ALIGNMENTS

RESULT 1
AI770311/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

AI770311
42 Mycobacterium anaerobic stationary phase library
smegmatis cDNA, mRNA sequence.
AI770311
AI770311.1
GI:6742680
EST.
Mycobacterium smegmatis.
Mycobacterium smegmatis
Bacteria: Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium.
1 (bases 1 to 1282)
Murugasu-Oei, B., Tay, A. and Dick, T.
Upregulation of stress response genes and ABC transporters in
anaerobic stationary-phase Mycobacterium smegmatis
Mol. Gen. Genet. 262 (4-5), 677-682 (1999)
20092472
Contact: Murugasu-Oei, B.
Mycobacterium Laboratory
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore 117609, Republic of Singapore
Tel: 65 874 3011
Fax: 65 779 1117
Email: mcbom@imcb.nus.edu.sg
Insert Length: 1282 Std Error: 0.00
Seq primer: T3 Forward: T7 Backward.

FEATURES
source
1. .1282
Location/Qualifiers

```
/organism="Mycobacterium smegmatis"
/strain="mc2155"
/db_xref="taxon:172"
/clone_lib="Mycobacterium anaerobic stationary phase
library"
/lab_host="E. coli XL1-Blue MRF"
/notes="Vector: Lambda ZAP II; Bacilli were disrupted using
a Mini bead beater (Biospec). RNA was isolated using the
RNeasy protocol (Qiagen). Purified RNA was subjected to 2
rounds of digestion with RNase-free DNase I (Promega).
DNase I was heat-inactivated at 75oC for 5 min. and
removed by using RNeasy columns followed by phenol
extraction and ethanol precipitation. The RNA
preparations were confirmed to be free of genomic DNA
contamination by carrying out PCR and RT-PCR using the
Access kit (Promega) and primers specific for the
histone-like protein gene hlp (Lee et al., 1998). cDNA
was synthesized using random hexamer primers (Promega)
and Stratagene's cDNA synthesis kit. cDNA fragments were
ligated into lambda ZAP II vector and packaged in vitro
using Stratagene reagents."
BASE COUNT      230 a 395 c 405 g 252 t
ORIGIN
Query Match      76.9%; Score 160; DB 9; Length 1282;
Best Local Similarity 85.6%; Pred. No. 5.6e-29;
Matches 178; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGTGTGCGCGGATACAAAGTCAACAGAGCTGG 60
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 236 TCAGAGAGAGCGCTACGACCTGGCCCGTGTGCGCGGTTTCAAGGTCAACAGAGCTGG 177
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 GCCTGAACACCAATATCGGATCACCAGGACGCTGACCGAAGAGACGCTGCTGCCCA 120
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 176 GCCTGAACCGGGCAAGCGGATCACCAGCTCGACGCTGACCGAGAGGACGCTGCTGCCGA 117
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 CCATCGAGTATCTGTCGCCCTGACAGAGGCGCCAGCAGATACCGTGGCGGCGGGG 180
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 116 CCATCGAGTACCTGTGCTGCTGCGACGAGGCTCAGACTTCGATGACCGCTCCCGGTGGCG 57
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 56 TCGAGGTTCCTCGAGGTTCGACGACAT 29
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 2
LOCUS      CNS0185F
DEFINITION Drosophila melanogaster genome survey sequence T7 end of RAC
      BACN01F12 of DrosHAC library from Drosophila melanogaster (fruit
      fly), genomic survey sequence.
ACCESSION  ALI08909
VERSION     ALI08909.1 GI:5629213
KEYWORDS   GSS.
SOURCE     Drosophila melanogaster.
      Drosophila melanogaster.
ORGANISM   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
      Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 1019)
AUTHORS   GSS.
TITLE     Direct Submission
JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
      BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
      - Web : www.genoscope.cns.fr)
COMMENT   Determination of this BAC-end sequence was carried out as part of a
      collaboration with the European Drosophila Genome Project (EDGP) -
      http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
      library (Dros BAC) was made by Alain Billaud at CEPH (Centre
      d'Etude du Polymorphisme Humain) with funding provided by a MRC
      project grant. The DNA was prepared from embryos by Alain Bucheton
      and Genevieve Payan. It has been constructed in the vector
      pBelOBAC11.
```

```
FEATURES
source      Location/Qualifiers
1..1019
      /organism="Drosophila melanogaster"
      /db_xref="taxon:7227"
      /clone="BACN01F12"
      /clone_lib="DrosBAC"
      /plasmid="pBelOBAC11"
      /notes="end : T7"
BASE COUNT  235 a 188 c 224 g 305 t 67 others
ORIGIN
Query Match      21.2%; Score 44; DB 17; Length 1019;
Best Local Similarity 51.0%; Pred. No. 0.67;
Matches 98; Conservative 2; Mismatches 92; Indels 0; Gaps 0;
QY 9 AAGUGCTACGACCTGGCCCGTGTGCGCGGATACAAGGTCAACAAGAGCTGGGCTGAAC 68
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 171 AAGCAAAAGATCGTGGGCAAMTTTGAAGCGGTTCAAGGTGAGCGAGCTGGAGGGGAGAK 230
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 69 ACCAATCATCGATCACCACGACGCGTACCAGGAAGACGCTGCTGCCGCCACCATCGAG 128
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 231 TCCGTCATCGAGCAGCGGCGGCTGCGCGGGAACGCGACCTGGATGCGCTCTTCCAG 290
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 129 TATCTGTCGCCCTGACAGAGGCGCCAGCATGACCGTGCCTGGCGGGGTCGAGGTG 188
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 291 GAGCTGGAGTSCGTCTCTGTTGCGAGGCGGACGACTCCGCGACCGACATGGACAGCATA 350
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 189 CCGGTGGAAACC 200
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 351 TCGGTGGGCTCC 362
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 3
LOCUS      AT405818
DEFINITION GH25953.3prime GH Drosophila melanogaster head pOT2 Drosophila
      melanogaster cDNA clone GH25953 5prime similar to AF001796: Krt95D
      F9gn0010501 PID:g2258302, mRNA sequence.
ACCESSION  AT405818
VERSION     AT405818.1 GI:4248905
KEYWORDS   EST.
SOURCE     fruit fly.
      Drosophila melanogaster
ORGANISM   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
      Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 630)
AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
      Lewis,S. and Rubin,G.M.
      BDGP/HHMI Drosophila EST Project
      Unpublished (2001)
      Contact: Stapleton, M.
      BDGP
      Lawrence Berkeley National Lab
      One Cyclotron Rd, Berkeley, CA 94720, USA
      Fax: 510 486 6798
      Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
      Plate: 259 row: E column: 5
      High quality sequence stop: 481.
FEATURES
source      Location/Qualifiers
1..630
      /organism="Drosophila melanogaster"
      /db_xref="taxon:7227"
      /clone="GH25953"
      /clone_lib="GH Drosophila melanogaster head pOT2"
      /sex="male and female"
      /dev_stage="adult"
      /lab_host="DH5 - alpha"
      /notes="Organ: head; Vector: pOT2; Site_1: EcoRI; Site_2:
      XhoI; Sized fractionated cDNAs were directly ligated into
      pOT2. Plasmid cDNA library."
BASE COUNT  143 a 173 c 203 g 111 t
ORIGIN
```

```
Query Match      20.8%; Score 43.2; DB 9; Length 630;
Best Local Similarity 51.6%; Pred. No. 0.99;
Matches 99; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 9 AAGCGCTACGACCTGCCCGTGTGCGCGATACAAAGGTCAACAAGAGCTGGCGCTGAAC 68
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 190 AAGCAAAAGATCGTGCGACATTTTGAAGCGGTTTCAAGGTGAGCGAGGAGCTGGAGGAG 249
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 69 ACCAATCATCCGATCACCACGACGACGTGACCGAAGAGAGCTGCTCGCCACCATCCAG 128
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 250 TCCGTCATCGAGGACGGCGCGCTCGCGGAGAACCGGACCTGGATGGCGCTTTCAG 309
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 129 TATCTGGTCCGCGTCCAGGCGCGAGCGACGATGACCGTGGCGGGCGGCTCGAGGTG 188
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 310 GAGCTGGAGTCCGTCTCTGTTGCGAGGCGACGACTCCGGACCGGACATGGACAGCATA 369
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 189 CCGGTGGAAACC 200
    ||||| | |
Db 370 TCGGTGGGCTCC 381

RESULT 5
BF007092      397 bp      mRNA      linear      EST 06-OCT-2000
LOCUS      147141 Amblyomma americanum adult Lambda zap Express Amblyomma
DEFINITION      americanum cDNA, mRNA sequence.
ACCESSION      BF007092
VERSION      BF007092.1 GI:10707367
KEYWORDS      EST.
SOURCE      Amblyomma americanum.
ORGANISM      Amblyomma americanum
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
Parasitiformes; Ixodida; Ixodidae; Amblyomma.
REFERENCE      1 (bases 1 to 397)
AUTHORS      Hill,C.A and Gutierrez,J.A.
TITLE      Analysis of the expressed genome of the lone star tick, Amblyomma
            americanum (Acari:Ixodidae) using an expressed sequence tag
            approach
JOURNAL      Microb. Comp. Genomics 5 (2), 89-101 (2000) In press
COMMENT      Contact: Hill CA
            Animal Science Discovery Research
            Elanco Animal Health, A Division of Eli Lilly and Company
            PO Box 708, 2001 West Main St., Greenfield, IN 46140, USA
            Tel: 317 277 0826
            Fax: 317 277 4522
            Email: HILL_CATHERINE_A@LILLY.COM.

FEATURES
    source
    1..397
    /organism="Amblyomma americanum"
    /db_xref="taxon:6943"
    /clone_lib="Amblyomma americanum adult Lambda zap Express"
    /sex="Male, Female"
    /dev_stage="Adult"
    /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT      124 a 110 c 124 g 36 t
ORIGIN
    Query Match      19.5%; Score 40.6; DB 12; Length 397;
    Best Local Similarity 55.2%; Pred. No. 4;
    Matches 79; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 35 CCGATACAGGTCAACAAGAGCTGGCGCTGAACACCAATCATCCGATCACCACGAC 94
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 223 CCAGGACGAGCTCATCAACAGCTCAACAGGAGAAGAAGCAGTTTCCAGGAGCAGA 282
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 95 GCTGACCGAAGAGAGCTGCTGCCACCATCGAGTATCTGTGTCGCCCTGCAGAGGCCA 154
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 283 GAAGACCGCCGAAGACCTCCAGGCCACCGAGGACAAGGTGAACCACTGAACAAAGTCAA 342
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 155 GGCCACGATGACCGTCCCGGGCG 177
    ||||| | | | | |
Db 343 GCCCAAGCTGGAGACAGCGCTCG 365

RESULT 6
BI995346
LOCUS      1031026E07.y2 C. reinhardtii CC-1690, Stress II (normalized),
DEFINITION      Lambda zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION      BI995346
VERSION      BI995346.1 GI:16430136
KEYWORDS      EST.

Query Match      20.8%; Score 43.2; DB 9; Length 662;
Best Local Similarity 51.6%; Pred. No. 1;
Matches 99; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 9 AAGCGCTACGACCTGCCCGTGTGCGCGATACAAAGGTCAACAAGAGCTGGCGCTGAAC 68
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 190 AAGCAAAAGATCGTGCGACATTTTGAAGCGGTTTCAAGGTGAGCGAGGAGCTGGAGGAG 249
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 69 ACCAATCATCCGATCACCACGACGACGTGACCGAAGAGAGCTGCTCGCCACCATCCAG 128
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 250 TCCGTCATCGAGGACGGCGCGCTCGCGGAGAACCGGACCTGGATGGCGCTTTCAG 309
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 129 TATCTGGTCCGCGTCCAGGCGCGAGCGACGATGACCGTGGCGGGCGGCTCGAGGTG 188
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 310 GAGCTGGAGTCCGTCTCTGTTGCGAGGCGACGACTCCGGACCGGACATGGACAGCATA 369
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 189 CCGGTGGAAACC 200
    ||||| | |
Db 370 TCGGTGGGCTCC 381

RESULT 4
A1405617      662 bp      mRNA      linear      EST 19-APR-2001
LOCUS      GH25702.5prime GH Drosophila melanogaster head pot2 Drosophila
DEFINITION      melanogaster cDNA clone GH25702 5prime similar to AF001796: Krt95D
            FBgn0010501 PID:g2258302, mRNA sequence.
ACCESSION      A1405617
VERSION      A1405617.1 GI:4248704
KEYWORDS      EST.
SOURCE      fruit fly.
ORGANISM      Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachyceta; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 662)
AUTHORS      Harvey,D., Broksstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
TITLE      BDGP/RHMI Drosophila EST Project
JOURNAL      Unpublished (2001)
COMMENT      Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 257 row: A column: 2
            High quality sequence stop: 564.

FEATURES
    source
    1..662
    /organism="Drosophila melanogaster"
    /db_xref="taxon:7227"
    /clone_lib="GH25702"
    /clone_lib="GH Drosophila melanogaster head pot2"
    /sex="male and female"
    /dev_stage="adult"
    /lab_host="rh5 - alpha"
    /note="Organ: head; Vector: pot2; site_1: EcoRI; Site_2:
            XhoI; Sized fractionated cDNAs were directly ligated into
            pot2. Plasmid cDNA library."

BASE COUNT      151 a 183 c 212 g 116 t
ORIGIN
    Query Match      20.8%; Score 43.2; DB 9; Length 662;
    Best Local Similarity 51.6%; Pred. No. 1;
    Matches 99; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 9 AAGCGCTACGACCTGCCCGTGTGCGCGATACAAAGGTCAACAAGAGCTGGCGCTGAAC 68
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 190 AAGCAAAAGATCGTGCGACATTTTGAAGCGGTTTCAAGGTGAGCGAGGAGCTGGAGGAG 249
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 69 ACCAATCATCCGATCACCACGACGACGTGACCGAAGAGAGCTGCTCGCCACCATCCAG 128
```


TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model.
JOURNAL Unicellular System for Analyzing Gene Function and Regulation in
COMMENT Vascular Plants. Project: 1030
 Unpublished (2002)
 Contact: Charles Hauser
 DCMB Box 91000
 Durham, NC 27708-1000
 Tel: 919 613 8159
 Fax: 919 613 8177
 Email: chauser@duke.edu.

FEATURES

SOURCE 1..666 /organism="Chlamydomonas reinhardtii" /strain="CC-1690 wild type ml+ 21gr" /db_xref="taxon:3055" /clone_lib="c. reinhardtii CC-1690, Deflagellation (normalized), Lambda Zap II" /note="Vector: pBluescript II SK-; Site.1: EcoRI; Site.2: XhoI; Deflagellation library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells which had been re-synthesizing flagella for 15, 30 and 60 min after being deflagellated by pH shock. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoRI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with EXassist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806." HASE COUNT 117 a 232 c 209 g 107 t 1 others ORIGIN	Query Match 18.8%; Score 39.2; DB 14; Length 666; Best Local Similarity 51.7%; Pred. No. 9.3; Matches 89; Conservative 0; Mismatches 83; Indels 0; Gaps 0; QY 6 GAGAAAGCGCTACGACCTGGCGCGGTGTGGCGGATACAAAGTCAACAAGAGCTGGGCGTG 65 Db 60 CAGAAAGTCGGCGCGCTGGGCGCGACCGCGTCCGCGCAAGTCCACACGCTGCTGGCGTG 119 QY 66 AACACCAATCATCGATCACCACACGACCTGACCGAAGAGACGTCGTGCGCCACCATC 125 Db 120 TACCGCATGTTTCGAGCTGGGCAAGGCGCGATCTTGGTCGACGCGGTGGACATCGCCACC 179 QY 126 GAGTATCTGCTCCCGCTCCACGAGGCGGACGCCACGATGACCGTCCGCGGCG 177 Db 180 CTGTGCGTCAAGCGCGCTCGCGACCGCGCTGTCCATCATTCGCCAGGAGCCCG 231
---	--

RESULT 9
 B0817941
LOCUS B0817941
DEFINITION B0817941 681 bp mRNA linear EST 01-AUG-2002
 Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
ACCESSION B0817941
VERSION B0817941.1 GI:22067958
KEYWORDS EST.
SOURCE Chlamydomonas reinhardtii.
ORGANISM Chlamydomonas reinhardtii
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadaceae; Chlamydomonas.

REFERENCE 1 (bases 1 to 681)
AUTHORS Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre,
 P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
 Unicellular System for Analyzing Gene Function and Regulation in
 Vascular Plants. Project: 1030
JOURNAL Unpublished (2002)
COMMENT Contact: Charles Hauser
 DCMB Box 91000
 Durham, NC 27708-1000

```

BASE COUNT      78 a   137 c   181 g   67 t   1 others
ORIGIN

Query Match      18.8%; Score 39; DB 13; Length 464;
Best Local Similarity 52.8%; Pred. No. 10;
Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 47 CAACAAGAGCTGGCGCTGAACACGATCATCCGATCACCACGACGCGTGCACCGAAGA 106
    ||| | ||| ||| ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
DB 124 CAAGAGCGCTGCTCTGGGTCCAAACCGGTGCACAAGACGCTGCCGGTCTCTCTCCA 188

QY 107 AGACGTCGTCGCCACCATCGAGTATCTGGTCCGCTGCACGAGGCGCAGCGATGAC 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 189 GCGCGCGCGCCCAATAACAGTCCCAAGTCTCATCTGCAGTATACATCGACGAGGTGTGGC 248

QY 167 CQTCCGCGCGGGTTCGAGGTGCGCGTGGAAACGACGA 205
    || ||| || ||| || ||| || ||| || ||| || ||| || ||| || ||| ||
DB 249 GGGACGCGCGGACGCTGCTCGCGCGACCCCTACGA 287

RESULT 11
LOCUS      BF008145          540 bp      mRNA      linear      EST 06-OCT-2000
DEFINITION 1682830 Amblyomma americanum adult Lambda Zap Express Amblyomma
ACCESSION  BF008145
VERSION     BF008145
KEYWORDS   EST.
SOURCE     Amblyomma americanum.
ORGANISM   Amblyomma americanum.

REFERENCE:
AUTHORS    Hill,C.A and Gutierrez,J.A.
TITLE      Analysis of the expressed genome of the lone star tick, Amblyomma
            americanum (Acari:Ixodidae) using an expressed sequence tag
            approach
JOURNAL    Microb. Comp. Genomics 5 (2), 89-101 (2000) In press
COMMENT    Contact: Hill CA
            Animal Science Discovery Research
            Elanco Animal Health, A Division of Eli Lilly and Company
            PO Box 708, 2001 West Main St., Greenfield, IN 46140, USA
            Tel: 317 277 0826
            Fax: 317 277 4522
            Email: HILL_CATHERINE_A@LILLY.COM.
FEATURES
  source
    1..540
    /organism="Amblyomma americanum"
    /db_xref="taxon:6943"
    /clone_lib="Amblyomma americanum adult Lambda Zap Express"
    /sex="Male, Female"
    /dev_stage="Adult"
    /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT      164 a   149 c   175 g   52 t
ORIGIN

Query Match      18.8%; Score 39; DB 12; Length 540;
Best Local Similarity 54.5%; Pred. No. 10;
Matches 78; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 35 CCGATACAAGGTCAACAAGAGCTGGCGCTGAACACCAATCATCCGATCACACGAGGAC 94
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 CCAGGACGAGTCAATCAACAAAGCTCAACAAGAGAGAGACGATTGCAGGAGGACCA 421

QY 95 GCTGACCGAAGACAGCTCTCGCCACCATCGAGTATCTGTCGCTCGACGAGGCGCA 154
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 422 GAAGACCGCGAGACCTCCAGCAACCGAGAGAGAGGTAACCACTGAACAAAGGTCAA 481

QY 155 GGCCACCATGACCGTCGCGGGCG 177
    |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 482 GGCCAAGCTGGACAGACGCTCG 504

```

```

RESULT 12
LOCUS      BG321153          710 bp      mRNA      linear      EST 27-FEB-2001
DEFINITION cDNA clone Zm04_05f06, mRNA sequence.
ACCESSION  BG321153
VERSION     BG321153
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 710)
AUTHORS    Singh,J.A., Wakui,K., Couroux,P., De Moors,A., Harris,L.J., Hattori
            ,J.I., Ouellet,T., Robert,L.S., Sprott,D. and Tinker,N.A.
            Expressed Sequence Tags from Cold-Stressed Maize Seedlings
            Unpublished (2001)
            Contact: Singh,J.A.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
            OC6, Canada
            Tel: (613) 759-1662
            Fax: (613) 759-1701
            Email: singhja@em.agr.ca.
            Location/Qualifiers
            1..710
            /organism="Zea mays"
            /cultivar="CO328"
            /db_xref="taxon:4577"
            /clone_lib="Zm04_05f06"
            /tissue_type="Leaf, crown"
            /note="Vector: Bluescript SK-/XhoI-EcoRI; Site_1: Eco RI;
            Site_2: Xho I; Lower temperature 50 C / hour from 22 to
            12oC; bring to 50 in 1 hour from 12oC. Leave at 50C 2 days
            , photoperiod 16 hours. Light intensity was 125 uE-l.
            Library prepared by in vivo mass excision from amplified
            library."

BASE COUNT      118 a   263 c   213 g   80 t   36 others
ORIGIN

Query Match      18.8%; Score 39; DB 12; Length 710;
Best Local Similarity 59.5%; Pred. No. 10;
Matches 66; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 45 GTCAACAAGAAGCTGGCGCTGAACACCAATCATCCGATCACACGACGCGTGACCGAA 104
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 388 GGCATCCAGGAMCTGCCAGGGTCCCATGGAGCAGATCGACGCCGCCAAGCTCACCAAG 447

QY 105 GAAGACGTCGTGCCACCATCGAGTATCTGTCGCTCGACGAGGCGCAG 155
    || ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 448 GAGCAGCTCGAGCTCYGCGCTCAAGTTGTCGCGCTCTTCACGCGCGCAAG 498

```

```

RESULT 13
LOCUS      BG836339          762 bp      mRNA      linear      EST 25-MAY-2001
DEFINITION Zm06_01a02_R
            Zm06_AAFc-ECORC_Fusarium-graminearum_inoculated_corn_eartip Zea
            mays cDNA clone Zm06_01a02, mRNA sequence.
ACCESSION  BG836339
VERSION     BG836339
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 762)
AUTHORS    Harris,L.J., Balcerzak,M., Allard,S., Saparno,A., Couroux,P., De
            Moors,A., Hattori,J.I., Ouellet,T., Robert,L.S., Singh,J.A, Sprott

```

TITLE D. and Tinker, N.A.
JOURNAL Expressed Sequence Tags from Maize Ear Tips 48 Hours after Silk
COMMENT Channel Inoculation with Fusarium graminearum
Unpublished (2001)
Contact: Harris, Linda J.
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1314
Fax: (613) 759-6566
Email: harrisljeem.agr.ca.

FEATURES Location/Qualifiers
1..762
/organism="Zea mays"
/cultivar="CO387"
/db_xref="taxon:4577"
/clone="Zm06_01a02"
/clone_lib="Zm06_AAFC_ECORC_Fusarium_graminearum_inoculate
d_corn_ear_tips"
/tissue_type="2.5 cm of top of unfertilized ear"
/dev_stage="6-7 days post-silk emergence"
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI;
Site_2: XhoI; Field-grown unpollinated maize ears were
silk channel-inoculated in the morning (~10 am) with 1 ml
of a Fusarium graminearum macroconidial suspension (500
000 spores/ml) and whole ears were collected and
immediately frozen in liquid nitrogen 48 hours later."
BASE COUNT 124 a 223 c 280 g 123 t 12 others
ORIGIN

Query Match 18.8%; Score 39; DB 12; Length 762;
Best Local Similarity 52.8%; Pred. No. 11;
Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 47 CAACAAGAGCTGGGCTGCAACCAATCATCCGATCACACGACGACCTGACCGAAGA 106
DB 190 CAAGAGCGCGCTCTCTCGGGTCCAACCCGGTGCACAGAGCGTGGCGGTCTCTCCA 249
QY 107 AGACGTCTGCGGCACCATCGAGTATCTGTGCGCCCTGACAGAGGCCAGGCCACGATGAC 166
DB 250 CGCCGGCGCGCGCAAAACGAGTCCGAGGTATCTCTGCGATACATCGAGGTGTGGGC 309
QY 167 CTGCGCGCGCGGGTGGAGTCTGCGGTGGAACCGACGA 205
DB 310 GGGGACGGCGCGAGGTCTCCCGACGACCCCTACGA 348

RESULT 14
AY104676
LOCUS AY104676 996 bp mRNA linear HTC 25-MAY-2002
DEFINITION Zea mays PC0085803 mRNA sequence.
ACCESSION AY104676
VERSION AY104676.1 GI:21207754
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 996)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
REFERENCE 2 (bases 1 to 996)
AUTHORS Coe, E.C.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
LOCATION/Qualifiers
1..996
source

/organism="Zea mays"
/db_xref="MaizeDB:634819"
/db_xref="taxon:4577"
/clone="PC0085803"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
BASE COUNT 193 a 286 c 326 g 191 t
ORIGIN

Query Match 18.8%; Score 39; DB 11; Length 996;
Best Local Similarity 52.8%; Pred. No. 11;
Matches 84; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 47 CAACAAGAGCTGGGCTGAAACCAATCATCCGATCACACGACGACCTGACCGAAGA 106
DB 207 CAAGAGCGCGCTCTCTCGGGATCCAAACCCGGTGCACAAAGAGCGTGGCGGTCTCTCCA 266
QY 107 AGACGTCTGCGGCACCATCGAGTATCTGTGCGCCCTGACAGAGGCCAGGCCACGATGAC 166
DB 267 CGCCGGCGCGCGCAAAACGAGTCCGAGGTATCTCTGCGATACATCGAGGTGTGGGC 326
QY 167 CTGCGCGCGCGGGTGGAGTCTGCGGTGGAACCGACGA 205
DB 327 GGGGACGGCGCGAGGTCTCCCGACGACCCCTACGA 365

RESULT 15
AY109418
LOCUS AY109418 1856 bp mRNA linear HTC 25-MAY-2002
DEFINITION Zea mays CL2386_1 mRNA sequence.
ACCESSION AY109418
VERSION AY109418.1 GI:21213135
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1856)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S.,
Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
REFERENCE 2 (bases 1 to 1856)
AUTHORS Coe, E.C.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
LOCATION/Qualifiers
1..1856
source

FEATURES Location/Qualifiers
1..1856
/organism="Zea mays"
/db_xref="MaizeDB:630950"
/db_xref="taxon:4577"
/clone="CL2386_1"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
BASE COUNT 320 a 548 c 585 g 281 t 122 others
ORIGIN
Query Match 18.8%; Score 39; DB 11; Length 1856;
Source

Best Local Similarity 59.5%; Pred. No. 12:
Matches 66; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
QY 45 GTCAACAGAGCTGGGCTGAACACCAATCATTCGATCACCCAGCAGCGCTGACCGAA 104
Dh 140 GGCATCCAGGACCTGGCCAGGCTCCCATCGAGCAGATCGAGCGCGCCAGCTCACCAG 199
QY 105 GAAGAGCTGCTGCCACCATTCGATATCTGGTCCGCCCTGCACGAGGGCCAG 155
Db 200 CACGAGCTCGAGTCCGGCTCAAGTGGCTCGGCCCTCTCCACCGCGCCGCAAG 250

Search completed: November 13, 2002, 03:50:54
Job time : 1122.63 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

1M nucleic - nucleic search, using sw model
Run on: November 12, 2002, 16:50:46 ; Search time 681.89 Seconds
(without alignments)
9517.553 Million cell updates/sec

Title: US-09-697-123b-13
Perferr score: 223
Sequence: 1 tcaagagagaagcgctacgac.....ccgggtggaacacgacgacat 223

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	
GenEmbl.*	
1: gb_ba.*	
2: gb_htg.*	
3: gb_in.*	
4: gb_om.*	
5: gb_ov.*	
6: gb_pat.*	
7: gb_ph.*	
8: gb_pl.*	
9: gb_pr.*	
10: gb_ro.*	
11: gb_sts.*	
12: gb_sy.*	
13: gb_un.*	
14: gb_vi.*	
15: em_ba.*	
16: em_fun.*	
17: em_hum.*	
18: em_in.*	
19: em_mu.*	
20: em_om.*	
21: em_or.*	
22: em_ov.*	
23: em_pat.*	
24: em_ph.*	
25: em_pl.*	
26: em_ro.*	
27: em_sts.*	
28: em_un.*	
29: em_vi.*	
30: em_htg_hum.*	
31: em_htg_inv.*	
32: em_htg_other.*	
33: em_htg_mus.*	
34: em_htg_pln.*	
35: em_htg_rod.*	
36: em_htg_mam.*	
37: em_htg_vrt.*	
38: em_sy.*	
39: em_htgo_hum.*	
40: em_htgo_mus.*	
41: em_htgo_other.*	

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	123.8	55.5	610	1	MTU318818	AJ318818 Mycobacte
2	123.8	55.5	610	1	MTU318819	AJ318819 Mycobacte
3	123.8	55.5	616	1	MTU318813	AJ318813 Mycobacte
4	123.8	55.5	618	1	MTU318815	AJ318815 Mycobacte
5	123.8	55.5	618	1	MTU318817	AJ318817 Mycobacte
6	123.8	55.5	633	1	MTU318814	AJ318814 Mycobacte
7	123.8	55.5	637	1	MTU318816	AJ318816 Mycobacte
8	123.8	55.5	639	1	MTU318821	AJ318821 Mycobacte
9	123.8	55.5	970	6	I50706	Sequence 1
10	123.8	55.5	3534	6	AX111339	Sequence 1
11	123.8	55.5	3853	1	MTU12205	U12205 Mycobacteri
12	123.8	55.5	5084	1	MSGRPOB	L27989 Mycobacteri
13	123.8	55.5	19352	1	AE006964	AE006964 Mycobacte
14	123.8	55.5	19770	1	MTC1376	295972 Mycobacteri
15	122.8	55.1	3752	1	MSU24494	U24494 Mycobacteri
16	111	49.8	3447	6	AR067447	AR067447 Sequence
17	111	49.8	37617	1	MLB1790G	Z14314 M.leprae ge
18	111	49.8	348950	1	MLEPRTN7	AL583923 Mycobacte
19	88.4	39.6	3941	1	AF242549	AF242549 Amycolato
20	68.8	30.9	32923	1	SCD82	AL160431 Streptomy
21	62.8	28.2	3495	6	AX120631	AX120631 Sequence
22	62.8	28.2	328050	1	AP005275	AP005275 Corynebac
23	62.8	28.2	349980	6	AX127144	AX127144 Sequence
24	46.8	21.0	204050	1	AL646070	AL646070 Ralstonia
25	45.2	20.3	121544	2	AC092779	AC092779 Oryza sat
26	45.2	20.3	150446	2	AC097277	AC097277 Oryza sat
27	44.6	20.0	139467	1	D90905	D90905 Synecocyst
28	44.2	19.8	204050	1	AL646073	AL646073 Ralstonia
29	42.2	18.9	2164	1	SC2C3	AL391157 Streptomy
30	42	18.8	3849	6	AX058889	AX058889 Sequence
31	42	18.8	10809	1	AE002027	AE002027 Deinococc
32	42	18.8	12744	1	AF204401	AF204401 Streptomy
33	41.4	18.6	2771	1	PAU70470	U70470 Pseudomonas
34	41.4	18.6	12948	1	AE004589	AE004589 Pseudomon
35	41.4	18.6	188050	1	AL646072	AL646072 Ralstonia
36	41.2	18.5	3304	6	E03373	E03373 DNA sequenc
37	41.2	18.5	3306	6	I28835	I28835 Sequence 71
38	41	18.4	149777	2	AP003573	AP003573 Oryza sat
39	40.8	18.3	496	1	SBU72186	U72186 Streptomyce
40	40.8	18.3	599	1	AF130448	AF130448 Deinococc
41	40.8	18.3	7954	1	TAQ19223	Y19223 Thermus aqu
42	40.8	18.3	166753	8	AP002866	AP002866 Oryza sat
43	40.6	18.2	1057	1	AB085278	AB085278 Unculture
44	40.6	18.2	1057	1	AB085282	AB085282 Unculture
45	40.6	18.2	39207	1	SC4G1	AL391039 Streptomy

ALIGNMENTS

RESULT 1	MTU318818	610 bp	DNA	circular BCT 09-AUG-2002
LOCUS	MTU318818			
DEFINITION	Mycobacterium tuberculosis partial rpoB gene for RNA polymerase			
ACCESSION	AJ318818			
VERSION	AJ318818.1 GI:22208412			
KEYWORDS	RNA polymerase beta subunit; rpoB gene.			
SOURCE	Mycobacterium tuberculosis.			
ORGANISM	Mycobacterium tuberculosis.			
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
AUTHORS	Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium tuberculosis complex.			
	Herrera, L., Jimenez, M.S. and Saez, J.A.			

TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
gene and review
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 610)
AUTHORS Herrera, L.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source Location/Qualifiers
1..610
/organism="Mycobacterium tuberculosis"
/isolate="1415-97"
/db_xref="taxon:1773"
gene 1..610
/gene="rpoB"
CDS <1..>610
/codon_start=1
/product="RNA polymerase beta subunit"
/protein_id="CAC87035.1"
/db_xref="GI:22208413"
/translation="LDIYRKLRPGEPTTKESAQTLLLENLFFKEXRYDLARVGRYKVNK
KLGLHVGEPITSSLTTEEDVVATIEYLRLHEGOTMTVPGVEVPVETDDIDHFGNR
RLRTVCGELIQNIRVGMSEMRVVRWTTQDVEAITPOTLINIRPVVAAIKKEFFGTS
QPSQFMGONNPLSLGTHKRRLSALPGLSRERAGLEVRDHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match 55.5%; Score 123.8; DB 1; Length 610;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
OY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCGCTACAAAGTCAACAAGAGCTGG 60
DB 80 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCGCTACAAAGTCAACAAGAGCTGG 139
OY 61 GCCTGAACACCGATCATCGATCACCACACGACGCTGACCGAAGACGCTGCGGCCA 120
DB 140 GCCTGATGTCGGCAGGCCATCAGCTGTCGACGCTGACCGAAGACGCTGCGGCCA 199
OY 121 CCATCGAGTACCTGGTTCGCCTGCACACGCTCTCTCAGGCTGGCCAGGCCCGCTTATGA 180
DB 200 CCATCGAATATCTGCTCCGCTTGCACGAGG-----GTCAGACACCATGA 244
OY 181 CTGTCCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
DB 245 CCGTTCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 287
RESULT 2
MTU318819 610 bp DNA circular BCT 09-AUG-2002
LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
DEFINITION beta subunit, isolate 1417-97.
ACCESSION AJ318819
VERSION AJ318819.1 GI:22208414
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis.
REFERENCE 1
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
gene and review
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source Location/Qualifiers
1..616
/organism="Mycobacterium tuberculosis"
/isolate="1415-97"
/db_xref="taxon:1773"
gene 1..616
/gene="rpoB"
CDS <1..>616
/codon_start=1
/product="RNA polymerase beta subunit"
/protein_id="CAC87036.1"
/db_xref="GI:22208415"
/translation="LDIYRKLRPGEPTTKESAQTLLLENLFFKEXRYDLARVGRYKVNK
KLGLHVGEPITSSLTTEEDVVATIEYLRLHEGOTMTVPGVEVPVETDDIDHFGNR
RLRTVCGELIQNIRVGMSEMRVVRWTTQDVEAITPOTLINIRPVVAAIKKEFFGTS
QPSQFMGONNPLSLGTHKRRLSALPGLSRERAGLEVRDHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match 55.5%; Score 123.8; DB 1; Length 610;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
OY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCGCTACAAAGTCAACAAGAGCTGG 60
DB 80 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCGCTACAAAGTCAACAAGAGCTGG 139
OY 61 GCCTGAACACCGATCATCGATCACCACACGACGCTGACCGAAGACGCTGCGGCCA 120
DB 140 GCCTGATGTCGGCAGGCCATCAGCTGTCGACGCTGACCGAAGACGCTGCGGCCA 199
OY 121 CCATCGAGTACCTGGTTCGCCTGCACACGCTCTCTCAGGCTGGCCAGGCCCGCTTATGA 180
DB 200 CCATCGAATATCTGCTCCGCTTGCACGAGG-----GTCAGACACCATGA 244
OY 181 CTGTCCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
DB 245 CCGTTCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 287
RESULT 3
MTU318813 616 bp DNA circular BCT 09-AUG-2002
LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
DEFINITION beta subunit, isolate 1763-97.
ACCESSION AJ318813
VERSION AJ318813.1 GI:22208402
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis.
REFERENCE 1
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
gene and review
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source Location/Qualifiers
1..616
/organism="Mycobacterium tuberculosis"
/isolate="1415-97"
/db_xref="taxon:1773"
gene 1..616
/gene="rpoB"
CDS <1..>616
/codon_start=1
/product="RNA polymerase beta subunit"
/protein_id="CAC87036.1"
/db_xref="GI:22208415"
/translation="LDIYRKLRPGEPTTKESAQTLLLENLFFKEXRYDLARVGRYKVNK
KLGLHVGEPITSSLTTEEDVVATIEYLRLHEGOTMTVPGVEVPVETDDIDHFGNR
RLRTVCGELIQNIRVGMSEMRVVRWTTQDVEAITPOTLINIRPVVAAIKKEFFGTS
QPSQFMGONNPLSLGTHKRRLSALPGLSRERAGLEVRDHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match 55.5%; Score 123.8; DB 1; Length 610;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
OY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCGCTACAAAGTCAACAAGAGCTGG 60
DB 80 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCGCTACAAAGTCAACAAGAGCTGG 139
OY 61 GCCTGAACACCGATCATCGATCACCACACGACGCTGACCGAAGACGCTGCGGCCA 120
DB 140 GCCTGATGTCGGCAGGCCATCAGCTGTCGACGCTGACCGAAGACGCTGCGGCCA 199
OY 121 CCATCGAGTACCTGGTTCGCCTGCACACGCTCTCTCAGGCTGGCCAGGCCCGCTTATGA 180
DB 200 CCATCGAATATCTGCTCCGCTTGCACGAGG-----GTCAGACACCATGA 244
OY 181 CTGTCCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
DB 245 CCGTTCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 287

TITLE Direct Submission
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source Location/Qualifiers
1..610
/organism="Mycobacterium tuberculosis"
/isolate="1417-97"
/db_xref="taxon:1773"
gene 1..610
/gene="rpoB"
CDS <1..>610
/codon_start=1
/product="RNA polymerase beta subunit"
/protein_id="CAC87036.1"
/db_xref="GI:22208415"
/translation="LDIYRKLRPGEPTTKESAQTLLLENLFFKEXRYDLARVGRYKVNK
KLGLHVGEPITSSLTTEEDVVATIEYLRLHEGOTMTVPGVEVPVETDDIDHFGNR
RLRTVCGELIQNIRVGMSEMRVVRWTTQDVEAITPOTLINIRPVVAAIKKEFFGTS
QPSQFMGONNPLSLGTHKRRLSALPGLSRERAGLEVRDHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match 55.5%; Score 123.8; DB 1; Length 610;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
OY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCGCTACAAAGTCAACAAGAGCTGG 60
DB 80 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCGCTACAAAGTCAACAAGAGCTGG 139
OY 61 GCCTGAACACCGATCATCGATCACCACACGACGCTGACCGAAGACGCTGCGGCCA 120
DB 140 GCCTGATGTCGGCAGGCCATCAGCTGTCGACGCTGACCGAAGACGCTGCGGCCA 199
OY 121 CCATCGAGTACCTGGTTCGCCTGCACACGCTCTCTCAGGCTGGCCAGGCCCGCTTATGA 180
DB 200 CCATCGAATATCTGCTCCGCTTGCACGAGG-----GTCAGACACCATGA 244
OY 181 CTGTCCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
DB 245 CCGTTCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 287
RESULT 3
MTU318813 616 bp DNA circular BCT 09-AUG-2002
LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
DEFINITION beta subunit, isolate 1763-97.
ACCESSION AJ318813
VERSION AJ318813.1 GI:22208402
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis.
REFERENCE 1
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
gene and review
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source Location/Qualifiers
1..616
/organism="Mycobacterium tuberculosis"
/isolate="1417-97"
/db_xref="taxon:1773"
gene 1..616
/gene="rpoB"
CDS <1..>616
/codon_start=1
/product="RNA polymerase beta subunit"
/protein_id="CAC87036.1"
/db_xref="GI:22208415"
/translation="LDIYRKLRPGEPTTKESAQTLLLENLFFKEXRYDLARVGRYKVNK
KLGLHVGEPITSSLTTEEDVVATIEYLRLHEGOTMTVPGVEVPVETDDIDHFGNR
RLRTVCGELIQNIRVGMSEMRVVRWTTQDVEAITPOTLINIRPVVAAIKKEFFGTS
QPSQFMGONNPLSLGTHKRRLSALPGLSRERAGLEVRDHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match 55.5%; Score 123.8; DB 1; Length 610;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
OY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCGCTACAAAGTCAACAAGAGCTGG 60
DB 80 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCGCTACAAAGTCAACAAGAGCTGG 139
OY 61 GCCTGAACACCGATCATCGATCACCACACGACGCTGACCGAAGACGCTGCGGCCA 120
DB 140 GCCTGATGTCGGCAGGCCATCAGCTGTCGACGCTGACCGAAGACGCTGCGGCCA 199
OY 121 CCATCGAGTACCTGGTTCGCCTGCACACGCTCTCTCAGGCTGGCCAGGCCCGCTTATGA 180
DB 200 CCATCGAATATCTGCTCCGCTTGCACGAGG-----GTCAGACACCATGA 244
OY 181 CTGTCCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
DB 245 CCGTTCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 287
RESULT 3
MTU318813 616 bp DNA circular BCT 09-AUG-2002
LOCUS Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
DEFINITION beta subunit, isolate 1763-97.
ACCESSION AJ318813
VERSION AJ318813.1 GI:22208402
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis.
REFERENCE 1
AUTHORS Herrera, L., Jimenez, M.S. and Saez, J.A.
TITLE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
gene and review
JOURNAL Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES
source Location/Qualifiers
1..616
/organism="Mycobacterium tuberculosis"
/isolate="1417-97"
/db_xref="taxon:1773"
gene 1..616
/gene="rpoB"
CDS <1..>616
/codon_start=1
/product="RNA polymerase beta subunit"
/protein_id="CAC87036.1"
/db_xref="GI:22208415"
/translation="LDIYRKLRPGEPTTKESAQTLLLENLFFKEXRYDLARVGRYKVNK
KLGLHVGEPITSSLTTEEDVVATIEYLRLHEGOTMTVPGVEVPVETDDIDHFGNR
RLRTVCGELIQNIRVGMSEMRVVRWTTQDVEAITPOTLINIRPVVAAIKKEFFGTS
QPSQFMGONNPLSLGTHKRRLSALPGLSRERAGLEVRDHP"
BASE COUNT 122 a 191 c 202 g 95 t
ORIGIN
Query Match 55.5%; Score 123.8; DB 1; Length 610;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
OY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCGCTACAAAGTCAACAAGAGCTGG 60
DB 80 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCGCTACAAAGTCAACAAGAGCTGG 139
OY 61 GCCTGAACACCGATCATCGATCACCACACGACGCTGACCGAAGACGCTGCGGCCA 120
DB 140 GCCTGATGTCGGCAGGCCATCAGCTGTCGACGCTGACCGAAGACGCTGCGGCCA 199
OY 121 CCATCGAGTACCTGGTTCGCCTGCACACGCTCTCTCAGGCTGGCCAGGCCCGCTTATGA 180
DB 200 CCATCGAATATCTGCTCCGCTTGCACGAGG-----GTCAGACACCATGA 244
OY 181 CTGTCCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
DB 245 CCGTTCCGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 287

```
/organism="Mycobacterium tuberculosis"  
/isolate="1763-97"  
/db_xref="taxon:1773"  
1..615  
/gene="rpoB"  
<1..>615  
/gene="rpoB"  
/codon_start=1  
/transl_table=1  
/product="RNA polymerase beta subunit"  
/protein_id="CAC87030.1"  
/db_xref="GI:22208403"  
/translation="LDIVKLRPGEPPTKESAQTLLLENLFFKPKRYDLARVYKVNK  
KLGLHVGEPITSSLTLEDVVATIEYLRLHEGQTTMPVGGVEVPVETDDIDHFGNR  
RLRTVCELIONQIRVGMKMEVYKHTTQDVRAITPOTLINRPVVAALKEFFGTS  
QLSQIFMDQNNPLSLGTHKKRLALGPGLSRERAGLEVRDVP"
```

```
BASE COUNT      125 a   191 c   201 g   99 t  
ORIGIN  
Query Match      55.5%; Score 123.8; DB 1; Length 616;  
Best Local Similarity 76.7%; Pred. No. 8.7e-14;  
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGCTCGCGCTACAAAGTCAACGAAGCTGG 60  
|||||  
Db 80 TCAAGGAGAGCGCTACGACCTGGCCGCGCTCGCGCTACAAAGTCAACGAAGCTGG 139  
|||||  
QY 61 CCTCAACACCGATCATCCGATCACCACCGACGCTGACCGAAGAACGCTGCGGCA 120  
|||||  
Db 140 GGTGCATGTCGGCGGAGCCATCAGCTCTGTCGACGCTGACCGAAGAACGCTGCGGCA 199  
|||||  
QY 121 CCATCGACGTACCTGCTCCCTGCGACCGACGCTCTCAGGGTGGCCGCGCGTTATGA 180  
|||||  
Db 200 CCATCGAATATCTGCTCCGCTTGCACGAGG-----GTCAAGACCAAGATGA 244  
|||||
```

```
Query Match      55.5%; Score 123.8; DB 1; Length 616;  
Best Local Similarity 76.7%; Pred. No. 8.7e-14;  
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGCTCGCGCTACAAAGTCAACGAAGCTGG 60  
|||||  
Db 80 TCAAGGAGAGCGCTACGACCTGGCCGCGCTCGCGCTACAAAGTCAACGAAGCTGG 139  
|||||  
QY 61 CCTCAACACCGATCATCCGATCACCACCGACGCTGACCGAAGAACGCTGCGGCA 120  
|||||  
Db 140 GGTGCATGTCGGCGGAGCCATCAGCTCTGTCGACGCTGACCGAAGAACGCTGCGGCA 199  
|||||  
QY 121 CCATCGACGTACCTGCTCCCTGCGACCGACGCTCTCAGGGTGGCCGCGCGTTATGA 180  
|||||  
Db 200 CCATCGAATATCTGCTCCGCTTGCACGAGG-----GTCAAGACCAAGATGA 244  
|||||  
QY 181 CTGTCCCGCGGGTCGAGGTGCGGTTGGAACCGACGACAT 223  
|||||  
Db 245 CGCTTCCGCGCGGTCGAGGTGCGGTTGGAACCGACGACAT 287  
|||||
```

```
RESULT 4  
MTU318815  
LOCUS  
DEFINITION  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 2540-97.  
ACCESSION  
AJ318815  
VERSION  
AJ318815.1 GI:22208406  
KEYWORDS  
RNA polymerase beta subunit; rpoB gene.  
SOURCE  
Mycobacterium tuberculosis.  
ORGANISM  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.
```

```
REFERENCE  
AUTHORS  
TITLE  
Herrera, L., Jimenez, M. S. and Saez, J. A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 618)  
AUTHORS  
TITLE  
Herrera, L.  
Direct Submission  
JOURNAL  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN
```

```
FEATURES  
source  
1..618  
/organism="Mycobacterium tuberculosis"  
/isolate="2540-97"  
/db_xref="taxon:1773"  
1..618  
/gene="rpoB"  
<1..>618  
gene  
CDS  
1..618  
/organism="Mycobacterium tuberculosis"  
/isolate="2540-97"  
/db_xref="taxon:1773"  
1..618  
/gene="rpoB"  
<1..>618
```

```
/gene="rpoB"  
/codon_start=1  
/transl_table=1  
/product="RNA polymerase beta subunit"  
/protein_id="CAC87032.1"  
/db_xref="GI:22208407"  
/translation="TDEALLDIYRKLPRGPPTKESAQTLLLENLFFKPKRYDLAKVGR  
YKVNKLGILHVGEPITSSLTLEDVVATIEYLRLHEGQTTMPVGGVEVPVETDDID  
HFGNRRLTVCGLIONQIRVGMKMEVYKHTTQDVRAITPOTLINRPVVAALKE  
FFGTSQLSQFMDQNNPLSLGTHKKRLALGPGLSRERAGLEVRDVP"  
BASE COUNT      124 a   190 c   207 g   97 t  
ORIGIN
```

```
Query Match      55.5%; Score 123.8; DB 1; Length 618;  
Best Local Similarity 76.7%; Pred. No. 8.7e-14;  
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;  
QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCGCTCGCGCTACAAAGTCAACGAAGCTGG 60  
|||||  
Db 95 TCAAGGAGAGCGCTACGACCTGGCCGCGCTCGCGCTACAAAGTCAACGAAGCTGG 154  
|||||  
QY 61 CCTCAACACCGATCATCCGATCACCACCGACGCTGACCGAAGAACGCTGCGGCA 120  
|||||  
Db 155 GGCTGCATGTCGGCGGAGCCATCAGCTCTGTCGACGCTGACCGAAGAACGCTGCGGCA 214  
|||||  
QY 121 CCATCGACGTACCTGCTCCCTGCGACCGACGCTCTCAGGGTGGCCGCGCGTTATGA 180  
|||||  
Db 215 CCATCGAATATCTGCTCCGCTTGCACGAGG-----GTCAAGACCAAGATGA 259  
|||||  
QY 181 CTGTCCCGCGGGTCGAGGTGCGGTTGGAACCGACGACAT 223  
|||||  
Db 260 CGTTCGCGCGGCTCGAGGTGCGGTTGGAACCGACGACAT 302  
|||||
```

```
RESULT 5  
MTU318817  
LOCUS  
DEFINITION  
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase  
beta subunit, isolate 2348-98.  
ACCESSION  
AJ318817  
VERSION  
AJ318817.1 GI:22208410  
KEYWORDS  
RNA polymerase beta subunit; rpoB gene.  
SOURCE  
Mycobacterium tuberculosis.  
ORGANISM  
Mycobacterium tuberculosis.  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium  
tuberculosis complex.
```

```
REFERENCE  
AUTHORS  
TITLE  
Herrera, L., Jimenez, M. S. and Saez, J. A.  
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis  
isolated in Spain (1996-2001). Description of new alleles into rpoB  
gene and review  
JOURNAL  
Unpublished  
REFERENCE  
2 (bases 1 to 618)  
AUTHORS  
TITLE  
Herrera, L.  
Direct Submission  
JOURNAL  
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro  
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,  
Majadahonda. Madrid. 28220, SPAIN
```

```
FEATURES  
source  
1..618  
/organism="Mycobacterium tuberculosis"  
/isolate="2348-98"  
/db_xref="taxon:1773"  
1..618  
/gene="rpoB"  
<1..>618  
/gene="rpoB"  
/codon_start=1  
/transl_table=1  
/product="RNA polymerase beta subunit"  
/protein_id="CAC87034.1"  
/db_xref="GI:22208411"
```

/translation="TDALLDIYKLRPGEPPTKESAQTLLLENLFFKRYDLARVGR
YKVNKGLHVGEDITSSTLTEDVATIEYLRLHEGQTTMTVPGGVEVPVETDDID
HFGNRRRTVGLIQNIQIRVMSRMVRVREMTTQDVEAITPTQLINIRPVVAAIKE
FFGTSQSQFMGONNPLSLGTHKRLSALPGGLSRERAGLEVDRV"
BASE COUNT 124 a 192 c 207 g 95 t
ORIGIN

Query Match 55.5%; Score 123.8; DB 1; Length 618;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCGCTGCGGCTACAAGGTCAACAAGAGCTGG 60
DB 95 TCAAGGAGAGCGCTACGACCTGGCCCGCGCTGCGGCTACAAGGTCAACAAGAGCTGG 154
QY 61 GCCTGAACACCGATCATCCGATCACCAACGACGCTGACCGAAGAGACGCTGCGCCA 120
DB 155 GCGTCGATGTCGGCGAGCCCATCAGCTGTCGACGCTGACCGAAGAGACGCTGCGCCA 214
QY 121 CCATCGAGTACCTGGTTCGCTGACCAACGACGCTCTCAGGTTGGCCAGGCCCGCTTATGA 180
DB 215 CCATCGAATATCTGGTTCGCTGACGAGG-----GTGACAGCACGATGA 259

QY 181 CTGTCCCGGGGGTCTGAGGTGCGCGTGGAAACGACGACAT 223
DB 260 CCGTTCGGGGCGGCTGAGGTGCGCGTGGAAACGACGACAT 302

RESULT 6
MTU318814
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .633
/organism="Mycobacterium tuberculosis"
/isolate="1058-97"
/db_xref="taxon:1773"
1. .633
/gene="rpoB"
/gene="rpoB"
/gene="rpoB"
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87031.1"
/db_xref="GI:22208405"
/translation="DNTVGTDEALLDIYKLRPGESPTKESAQTLLLENLFFKRYDL
ARVGRYKVNKGLHVGEDITSSTLTEDVATIEYLRLHEGQTTMTVPGGVEVPVE
TDDIDHFGNRRRTVGLIQNIQIRVMSRMVRVREMTTQDVEAITPTQLINIRPVV
AAIKEFFGTSQSQFMGONNPLSLGTHKRLSALPGGLSRERAGLEVDRV"
BASE COUNT 129 a 195 c 210 g 99 t
ORIGIN

Query Match 55.5%; Score 123.8; DB 1; Length 633;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;

QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCGCTGCGGCTACAAGGTCAACAAGAGCTGG 60
DB 110 TCAAGGAGAGCGCTACGACCTGGCCCGCGCTGCGGCTACAAGGTCAACAAGAGCTGG 169
QY 61 GCCTGAACACCGATCATCCGATCACCAACGACGCTGACCGAAGAGACGCTGCTGCCCA 120
DB 170 GCGTCGATGTCGGCGAGCCCATCAGCTGTCGACGCTGACCGAAGAGAGCTGCTGCCCA 229
QY 121 CCATCGAGTACCTGGTTCGCTGACCAACGACGCTCTCAGGTTGGCCAGGCCCGCTTATGA 180
DB 230 CCATCGAATATCTGGTTCGCTGACGAGG-----GTGACAGCACGATGA 274
QY 181 CTGTCCCGGGGGTCTGAGGTGCGCGTGGAAACGACGACAT 223
DB 275 CCGTTCGGGGCGGCTGAGGTGCGCGTGGAAACGACGACAT 317

RESULT 7
MTU318816
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Location/Qualifiers
1. .637
/organism="Mycobacterium tuberculosis"
/isolate="1255-98"
/db_xref="taxon:1773"
1. .637
/gene="rpoB"
/gene="rpoB"
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87033.1"
/db_xref="GI:22208409"
/translation="VGTDALLDIYKLRPGEPPTKESAQTLLLENLFFKRYDLARY
GRYKVNKGLHVGEPITSSTLTEDVATIEYLRLHEGQTTMTVPGGVEVPVETDD
IDHFGNRRRTVGLIQNIQIRVMSRMVRVREMTTQDVEAITPTQLINIRPVVAAI
KEFFGTSQSQFMGONNPLSLGTHKRRFALPGGLSRERAGLEVDRVHPSH"
BASE COUNT 128 a 198 c 210 g 101 t
ORIGIN

MTU318816
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1255-98.
AJ318816
AJ318816.1 GI:22208408
RNA polymerase beta subunit; rpoB gene.
Mycobacterium tuberculosis.
Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
Unpublished
2 (bases 1 to 637)
Herrera, L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN

MTU318816
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1255-98.
AJ318816
AJ318816.1 GI:22208408
RNA polymerase beta subunit; rpoB gene.
Mycobacterium tuberculosis.
Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
Unpublished
2 (bases 1 to 637)
Herrera, L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN

MTU318816
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1255-98.
AJ318816
AJ318816.1 GI:22208408
RNA polymerase beta subunit; rpoB gene.
Mycobacterium tuberculosis.
Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
Unpublished
2 (bases 1 to 637)
Herrera, L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN

MTU318816
Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1255-98.
AJ318816
AJ318816.1 GI:22208408
RNA polymerase beta subunit; rpoB gene.
Mycobacterium tuberculosis.
Mycobacterium tuberculosis.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
Herrera, L., Jimenez, M.S. and Saez, J.A.
Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
isolated in Spain (1996-2001). Description of new alleles into rpoB
gene and review
Unpublished
2 (bases 1 to 637)
Herrera, L.
Direct Submission
Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
Nacional Microbiologia., Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda, Madrid. 28220, SPAIN


```

101 TCAAGGAGAGCGCTAGGACCTGGCCCGCTCGTTCGCTATAGGTCAACAGAAAGCTCG 160
|||||
61 GCCTGAACACCGATATCCGATCACCACACGACGCTGACCGAAGAGACGTCGTGGCCA 120
|||||
161 GCGTGCATGTCGCGGAGCGCATCACCTCGTGCAGCTGACCGAAGAAGACGTCGTGGCCA 220
|||||
121 CCAATCGAGTACCTGGTTCCCTGCGACACGCGCTCTCAGGGTGGCCAGGCCCGCTTATGA 180
|||||
166 CCATCGAATATCTGGTCCGCTTCGACGAGG-----GTCAGACCAACGATGA 265

181 CTCTCCCGCGGGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
|||||
266 CACTTCCGGCGCGCTGAGGTGCGCGTGGAAACCGACGACAT 308

RESULT 8
LOCUS MTU318821 639 bp DNA circular BCT 09-AUG-2002
DEFINITION Mycobacterium tuberculosis partial rpoB gene for RNA polymerase
beta subunit, isolate 1071-98.
ACCESSION AJ318821
VERSION AJ318821.1 GI:22208418
KEYWORDS RNA polymerase beta subunit; rpoB gene.
SOURCE Mycobacterium tuberculosis.
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Mycobacteriaceae; Mycobacterium; Mycobacterium
tuberculosis complex.
1 Herrera,L., Jimenez,M.S. and Saez,J.A.
REFERENCE Molecular analysis of rifampin-resistant Mycobacterium tuberculosis
AUTHORS isolated in Spain (1996-2001). Description of new alleles into rpoB
TITLE gene and review
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 639)
AUTHORS Herrera,L.
JOURNAL Direct Submission
TITLE Submitted (07-AUG-2001) Herrera L., Servicio Bacteriologia, Centro
National Microbiologia, Ctra. Majadahonda-Pozuelo, Km 2.5,
Majadahonda. Madrid. 28220, SPAIN
FEATURES Location/Qualifiers
source
1..639
/organism="Mycobacterium tuberculosis"
/isolate="1071-98"
/db_xref="taxon:1773"
gene 1..639
/gene="rpoB"
CDS <1..>639
/gene="rpoB"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta subunit"
/protein_id="CAC87038.1"
/db_xref="GI:22208419"
/translation="RRTDEALLDIYVKRLPGEPPTKESAOTLLENLFKEKRYDLARV
GRYKNKGLGIVGEPITSSTLTEDVVATIEYLVRHGEQTTWVPGVEVPVETDD
IDHFCNRLRTVGLGELIQNIIVGMSRMRVYREKWTQDVAITPQILINRPAVAI
KEFTGSLSQFMVQNPPLSLGTLQKRRLSALGPGLSRERAGLEVDRVHPSHS"
BASE COUNT 126 a 202 c 212 g 99 t
ORIGIN
Query Match 55.5%; Score 123.8; DB 1; Length 639;
Best Local Similarity 76.7%; Pred. No. 8.7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;

QY 1 TCAAGGAGAGCGCTAGGACCTGGCCCGCTCGTTCGCTATAGGTCAACAGAAAGCTCG 60
|||||
DB 1 TCAAGGAGAGCGCTAGGACCTGGCCCGCTCGTTCGCTATAGGTCAACAGAAAGCTCG 85
|||||
QY 61 GCCTGAACACCGATATCCGATCACCACACGACGCTGACCGAAGAGACGTCGTGGCCA 120
|||||
DB 66 GCGTGCATGTCGCGGAGCGCCATCACGTCGTCGACGCTACCGAAGAGACGTCGTGGCCA 145
|||||
QY 121 CCAATCGAGTACCTGGTTCCCTGCGACACGCGCTCTCAGGGTGGCCAGGCCCGCTTATGA 180
|||||
DB 146 CCATCGAATATCTGGTCCGCTTCGACGAGG-----GTCAGACCAACGATGA 190
|||||
QY 181 CTGTCCCGCGGGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
|||||
DB 191 CCGTTCGGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 233
|||||

RESULT 10
LOCUS AX111339 3534 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 2072 from Patent W00123604.
ACCESSION AX111339
VERSION AX111339.1 GI:13927631
KEYWORDS
SOURCE Mycobacterium tuberculosis.
ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 3534)
REFERENCE Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
AUTHORS Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: W0 0123604-A 2072 05-APR-2001;
FEATURES Infectio Diagnostic (I.D.I.) INC. (CA)
Location/Qualifiers

```

```

QY 121 CCATCGAGTACCTGGTTCCCTGCGACACGCGCTCTCAGGGTGGCCAGGCCCGCTTATGA 180
|||||
DB 221 CCATCGAATATCTGGTCCGCTTCGACGAGG-----GTCAGACCAACGATGA 265
|||||
QY 181 CTGTCCCGCGGGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
|||||
DB 266 CCGTTCGGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 308
|||||

RESULT 9
LOCUS 150706 970 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 1 from patent US 5643723.
ACCESSION 150706
VERSION 150706.1 GI:2472409
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 970)
AUTHORS Persing,D.H., Hunt,J.J., Young,K.K.Y., Felmler,T.A., Roberts,G.D.
and Whelan,A.Christian.
TITLE Detection of a genetic locus encoding resistance to rifampin in
mycobacterial cultures and in clinical specimens
JOURNAL Patent: US 5643723-A 1 01-JUL-1997;
FEATURES Location/Qualifiers
source
1..970
/organism="unknown"
BASE COUNT 182 a 302 c 330 g 156 t
ORIGIN
Query Match 55.5%; Score 123.8; DB 6; Length 970;
Best Local Similarity 76.7%; Pred. No. 8.3e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;

QY 1 TCAAGGAGAGCGCTAGGACCTGGCCCGCTCGTTCGCTATAGGTCAACAGAAAGCTCG 60
|||||
DB 26 TCAAGGAGAGCGCTAGGACCTGGCCCGCTCGTTCGCTATAGGTCAACAGAAAGCTCG 85
|||||
QY 61 GCCTGAACACCGATATCCGATCACCACACGACGCTGACCGAAGAGACGTCGTGGCCA 120
|||||
DB 86 GCGTGCATGTCGCGGAGCGCCATCACGTCGTCGACGCTACCGAAGAGACGTCGTGGCCA 145
|||||
QY 121 CCATCGAGTACCTGGTTCCCTGCGACACGCGCTCTCAGGGTGGCCAGGCCCGCTTATGA 180
|||||
DB 146 CCATCGAATATCTGGTCCGCTTCGACGAGG-----GTCAGACCAACGATGA 190
|||||
QY 181 CTGTCCCGCGGGGTCGAGGTGCGCGTGGAAACCGACGACAT 223
|||||
DB 191 CCGTTCGGCGCGGTCGAGGTGCGCGTGGAAACCGACGACAT 233
|||||

RESULT 10
LOCUS AX111339 3534 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 2072 from Patent W00123604.
ACCESSION AX111339
VERSION AX111339.1 GI:13927631
KEYWORDS
SOURCE Mycobacterium tuberculosis.
ORGANISM Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 3534)
REFERENCE Bergeron,M.G., Boissinot,M., Huletsky,A., m Nard,C., Ouellette,M.,
AUTHORS Picard,F.J. and Roy,P.H.
TITLE Highly conserved genes and their use to generate probes and primers
for detection of microorganisms
JOURNAL Patent: W0 0123604-A 2072 05-APR-2001;
FEATURES Infectio Diagnostic (I.D.I.) INC. (CA)
Location/Qualifiers

```


REPLEVPGLLDVOTDSFEWLLIGSPWRRESAAERGDVNPVPGGLEEVLYELSPEDFSGS
NSLSFSDPRFDVVKAPVDECKDMTYAAPLVTAEPFINNTGTGKSTQVFMGDFPMH
TEKGTGFIINGTERVYVVSQIVKSPGVYFDETIDKSTDKTLHSVKVIPSRSQAWLEDFDVK
RDTGVRIIDRRKRQPVTVLLKALGWTSEQIVERFGFSEIMRSTLEKONTGVTGDEALLD
YRKLRLPEGPTKESAQTLLENLFFKRYDLARVGRYKVNKTKGLHVGEPITSTLT
EDDVVATIEYVRLHEGQTTMTVPGGVPEVDETDDHFGNRRRLRTVCGELQNOIRVG
NSMRVVRERMTTODVEAITPOTLINIRPVAAIKFEFGTSQSOQNPNPLSGLT
HKRRLSALPGSLSRERAGLEVROVHPSHYGRMCPLETPEGPNIGLIGLSVYARVNP
FCFLETPTPKVVDVYDEIVYLTADEDRHVYQAQNSPIEDAGGRFVPEVRLVRKAG
EVEYVPSSEVDYMDVSPROMVSVATAMIPFLEHDDANRALMGANMOQAAPLVLRSEAP
LVGTGMELRAAIDAATSSQESGVIEEVSADYITVMHNDGTRRTYRMKRFARSNGHTC
ANOCPIVDAGROEAGVITADGCTDDGEMALGNLLVAIMPMECHNYEDAIIILSNRL
DEEDVLTSTHIEEHIIDARDTKLGAEETTRDIPNTSDEVLDLDERGIVRIGAEVRDG
DILVCKVTPKGETELTPEERLLRAIFGEKAREVRDTSKVPHGSEKGVICIRVFSRED
EDELPAQVNLVRYVVAQRKISDGDKLGRHNGKVGIGKILPVDEMPFLADGTPVDI
LINTHGVPRMNIIGIILETHLGMCAHSGKVDAAKGVDPDWAARLPDELEHAHAIVS
TPVFGAQAELQGLISCTLPNRDGDVLVDADGKAMLFDRSGEPFPYPTVYAGMYIM
KLHLVDDKIHARSTGYSMTIQPPLGKAQFGGORGEMECWAMOYGAAYTLQELL
TIKSUDTVGRVKVYEAIVKGENIPEPGIPESFKVLLKELOSCLLNVEVLSDDGAALIEL
REGEDEDLERAAANLGINLSRNESASFEDLA"

gene
4641.. 5084
/gene="rpoC"
CDS
4641.. >5084
/gene="rpoC"
/codon_start=1
/transl_table=11
/product="RNA polymerase beta'-subunit"
/protein_id="AAA21417.1"
/db_xref="GI:537608"
/translation="MLDVNFFDELRLIGLATAEDIROMSYGEVKKPETINRYTLKPEKD
GLFCEKIFGPTRDWECYCGYKRVFRFGIICERCGEVTRAKVRERMGHIELAAPVT
HIWFKGVP SRLGYLLDLAPKDLKIIYFAAYVITSVDEMRHNEL"

BASE COUNT 969 a 1534 c 1691 g 890 t
ORIGIN
Query Match 55.5% Score 123.8; DB 1; Length 5084;
Best Local Similarity 76.7% Pred. No. 7e-14;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
QY 1 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCCGCTACAGGTCAACAAGAGCTGG 60
Db 1966 TCAAGGAGAGCGCTACGACCTGGCCCGCGTGGCCGCTATAGGTCAACAAGAGCTGG 2025
QY 61 GCGTCAACGATCATCCGATCACCACCGACCGCTGACCGAAGAGCGTGGTGGCCA 120
Db 2026 GCGTCAATCGCGGAGCCATCAGCTGCTCAGCGTCAACGAGAGCGTGGTGGCCA 2085
QY 121 CCATCAGTACCTGGTTCGCCCTGCACCGCCCTCTCAAGGTGGCCAGGCCCGCTTATGA 180
Db 2086 CCATCAATATCTGGTCCGCTTGCACGAGG-----GTCAGACACCATGA 2130
QY 181 CTGTCCCGCGGCTGCGAGGTGCGCGTGAACCGACGACAT 223
Db 2131 CCGTTCGCGGCGGCTGCGAGGTGCGCGTGAACCGACGACAT 2173

RESULT 13
LOCUS AE006964 19352 bp DNA linear BCT 27-APR-2001
DEFINITION Mycobacterium tuberculosis CDC1551, section 50 of 280 of the complete genome.
ACCESSION AE006964 AE000516
VERSION AE006964.1 GI:13880217
KEYWORDS
SOURCE Mycobacterium tuberculosis CDC1551.
ORGANISM Mycobacterium tuberculosis CDC1551
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacterineae; Mycobacteriaceae;
Mycobacterium; Mycobacterium tuberculosis complex.
1 (bases 1 to 19352)
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M.,

Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khour, H., Gill, J., Mikula, A. and Bishai, W.
Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains
Unpublished
2 (bases 1 to 19352)
Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khour, H., Gill, J., Mikula, A. and Bishai, W.
Direct Submission
Submitted (25-APR-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
1..19352
/organism="Mycobacterium tuberculosis CDC1551"
/strain="CDC1551"
/db_xref="taxon:8331"
/note="clinical strain"
163..3699
/gene="MT0695"
163..3699
/gene="MT0695"
note="similar to GB:L27989 GB:L05910 GB:U12205 SP:P47766
PID:149992; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA-directed RNA polymerase, beta subunit"
/protein_id="AAK44921.1"
/db_xref="GI:13880218"
/translation="MLEGCLADRSQKTAASPSRSPSSNNNSVPGAPNRVVSFAKL
REPLEVPGLLDVOTDSFEWLLIGSPWRRESAAERGDVNPVPGGLEEVLYELSPEDFSGS
MSLSFSDPRFDVVKAPVDECKDMTYAAPLVTAEPFINNTGTGKSTQVFMGDFPMH
TEKGTGFIINGTERVYVVSQIVKSPGVYFDETIDKSTDKTLHSVKVIPSRSQAWLEDFDVK
RDTGVRIIDRRKRQPVTVLLKALGWTSEQIVERFGFSEIMRSTLEKONTGVTGDEALLD
YRKLRLPEGPTKESAQTLLENLFFKRYDLARVGRYKVNKTKGLHVGEPITSTLT
EDDVVATIEYVRLHEGQTTMTVPGGVPEVDETDDHFGNRRRLRTVCGELQNOIRVG
NSMRVVRERMTTODVEAITPOTLINIRPVAAIKFEFGTSQSOQNPNPLSGLT
HKRRLSALPGSLSRERAGLEVROVHPSHYGRMCPLETPEGPNIGLIGLSVYARVNP
FCFLETPTPKVVDVYDEIVYLTADEDRHVYQAQNSPIEDAGGRFVPEVRLVRKAG
EVEYVPSSEVDYMDVSPROMVSVATAMIPFLEHDDANRALMGANMOQAAPLVLRSEAP
LVGTGMELRAAIDAATSSQESGVIEEVSADYITVMHNDGTRRTYRMKRFARSNGHTC
ANOCPIVDAGROEAGVITADGCTDDGEMALGNLLVAIMPMECHNYEDAIIILSNRL
DEEDVLTSTHIEEHIIDARDTKLGAEETTRDIPNTSDEVLDLDERGIVRIGAEVRDG
DILVCKVTPKGETELTPEERLLRAIFGEKAREVRDTSKVPHGSEKGVICIRVFSRED
EDELPAQVNLVRYVVAQRKISDGDKLGRHNGKVGIGKILPVDEMPFLADGTPVDI
LINTHGVPRMNIIGIILETHLGMCAHSGKVDAAKGVDPDWAARLPDELEHAHAIVS
TPVFGAQAELQGLISCTLPNRDGDVLVDADGKAMLFDRSGEPFPYPTVYAGMYIM
KLHLVDDKIHARSTGYSMTIQPPLGKAQFGGORGEMECWAMOYGAAYTLQELL
TIKSUDTVGRVKVYEAIVKGENIPEPGIPESFKVLLKELOSCLLNVEVLSDDGAALIEL
REGEDEDLERAAANLGINLSRNESASFEDLA"

REFERENCE
AUTHORS

1 (bases 1 to 19770)
 Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C.,
 Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry III, C. E.,
 Tekalaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T.,
 Connor, R., Davies, K., Devlin, K., Feltwell, T., Gentles, S.,
 Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J.,
 Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M. A.,
 Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, S.,
 Squares, S., Szares, R., Sulston, J. E., Taylor, K., Whitehead, S. and
 Barrell, B. G.
 Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence
 Nature 393 (6685), 537-544 (1998)
 98295987
 9634230
 2 (bases 1 to 19770)
 Parkhill, J.
 Direct Submission
 Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium
 tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
 Trust Genome Campus, Hinxton, Cambridge CB10 1SA Unite de Genetique
 Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
 75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
 On Jun 27, 1998 this sequence version replaced gi:2143285.

COMMENT

Details of M. tuberculosis sequencing at the Sanger Centre are
 available on the World Wide Web.
 (URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have
 been renumbered from the original cosmid submissions but the old
 gene designations are in brackets after the new gene numbers.
 Gene prediction was based on a Hidden Markov Model of TB genes
 implemented in TBParse (Krogh) supplemented with visual inspection
 of positional base preference in codons, especially where there is
 an increase in the observed/expected third position G + C.
 CAUTION: In some cases we may not have predicted the correct
 initiation codon. Where possible we choose an initiation codon
 (atg, gtg, or ttg) which is preceded by an upstream ribosome
 binding site sequence (optimally 5-13bp before the initiation
 codon). If this cannot be identified we choose the most upstream
 initiation codon.

FEATURES
Location/Qualifiers

1..19770
 /organism="Mycobacterium tuberculosis H37Rv"
 /strain="H37Rv"
 /db_xref="taxon:83332"
 /clone="I376"
 57..61
 /note="agga, possible rbs upstream of Rv0654"
 68..1573
 /gene="Rv0654"
 68..1573
 /gene="Rv0654"
 /note="Rv0654, (MTC1376.22), len: 501. unknown, FASTA
 score: 05353 LIGNOSTILBENE-ALPHA.BETA-DIOXYGENASE (485
 aa) opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity
 in 523 aa overlap). Also similar to M. tuberculosis protein
 MTCY21C12.07c (29.5% identity in 522 aa overlap)"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein Rv0654"
 /protein_id="CAB09380.1"
 /db_xref="GI:2143307"
 /db_xref="SPTREMBL:O06785"

RBS

gene

CDS

/note="Rv0654, (MTC1376.22), len: 501. unknown, FASTA
 score: 05353 LIGNOSTILBENE-ALPHA.BETA-DIOXYGENASE (485
 aa) opt:280 z-score: 330.1 E(): 2.3e-11, (28.5% identity
 in 523 aa overlap). Also similar to M. tuberculosis protein
 MTCY21C12.07c (29.5% identity in 522 aa overlap)"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein Rv0654"
 /protein_id="CAB09380.1"
 /db_xref="GI:2143307"
 /db_xref="SPTREMBL:O06785"
 /translation="MTTAAAESONPYLEGFLAPVSTEVATDLPVTCRIPEDHLDGRY
 LRNGPNPVAEVDPATYHWTGDAMVHGVALRDGRARWTRNRVTRPACVAALGEPISA
 RHPQRTGIEGGPNVLTTHAGRTLALVEAGVNVYELTDLTGPCDFDGLHGGYT
 AHQRDPHTGELHAVSYSFARGHRVQYSVIGTDGHARTVDIEVAGSPMMHFSLTND
 YVYIDLPVTFDPMPVPSVPRWLQRPALRVIOVLGRVRIPODIAALGNRMQGHSD
 RLPVWNPYSPARVYMPREGNEDVRWEDIEPCYVYHPLNAYSECRNCAEVLVDYV
 RYSRMFDRRPGCGDSRPSLDRTWTINLTCAVTAECDRDAQEFPRINLTGVGPHR
 FAYTVGIEGFLVGAGAAALSTPLYKQDCVTGSSVSLDPLDIGEMVFPVNPNSARAE
 DDGILMGYVHRGRDEGQLLLDQAOTLESTIATVHLQRPVPMGFHGNAPT"
 1585..2664

gene

CDS

/gene="Rv0655"
 1585..2664
 /gene="Rv0655"
 /note="Rv0655, (MTC1376.21), len: 359. abc transporter,
 FASTA score: YRBF_ECOLI_P45393 hypothetical abc
 transporter atp-binding (269 aa) opt: 644 z-score: 721.8
 E(): 3.4e-33 (38.5% identity in 244 aa overlap); contains
 PS00017 ATP/GTP-binding site motif A, PS00211 ABC
 transporters family signature, highly similar to M. leprae
 MKL_MYCLE_P30769 possible ribonucleotide transport atp-
 (347 aa) opt: 2021 z-score:2244.4 E(): 0. (92.2% identity
 in 335 aa overlap). Also similar to many other M.
 tuberculosis ABC transporters eg. MTCY253.24 (33.68
 identity in 241 aa overlap)"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein Rv0655"
 /protein_id="CAB09379.1"
 /db_xref="GI:2143306"
 /db_xref="SWISS-PROT:O06784"
 /translation="MRYSDSYHTTGRMQPRASTEGFPMGVSIEVNGLTSGFSGSRWIE
 DVTITPAGEVSVLLGPGSGTKSVFLSKLIGLLRPERGSIIDGTIIIECSAKELYEI
 RTLFGVLFDGALFGSMNLVYNTAPLREHTKKVESEIRDIWMKELINDVGLGGDEKFK
 PGEISGGMKRRAGLARALVDPQIILDEPDSGLDPVRTAVLSOLINDINAQIDATIL
 IVTNINARTVDPNMGMLFRKHLVMSGPREVILTSDEPVVRFQNLNGRRIGPICMSEE
 KDEATMAEEQALLDAGHHAGGVEIEGVPPQISATPGMPERRKAVRQAQVREMLHTL
 PKKAQAATLDDLEGTHKYAVHEIQG"
 1762..1785
 /gene="Rv0655"
 /note="PS00017 ATP/GTP-binding site motif A"
 2074..2118
 /gene="Rv0655"
 /note="PS00211 ABC transporters family signature"
 complement(3052..3435)
 /gene="Rv0656c"
 complement(3052..3435)
 /gene="Rv0656c"
 /note="Rv0656c, (MTC1376.20, unknown), len: 127 aa"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein Rv0656c"
 /protein_id="CAB09378.1"
 /db_xref="GI:2143305"
 /db_xref="SPTREMBL:O06783"
 /translation="MAAATTTGTHRGLELRAAQRAVGSCEPQRAEFCRSANNADEFDO
 MSRMFGDYVPDPVPKPSVWRWIDSAQHLRAGAVGALSVDLLICDTAAARGLVLLH
 DDADYELAERHLPDIRVRVVSADD"
 complement(3530..3685)
 /gene="Rv0657c"
 complement(3530..3685)
 /gene="Rv0657c"
 /note="Rv0657c, (MTC1376.19), unknown, len: 51 aa: similar
 to several other M. tuberculosis hypothetical proteins eg.
 yw08_MYCTU_Q10848 hypothetical 8.9 kd protein cy39.08c (80
 aa), fasta scores: opt: 107 z-score: 182.3 E(): 0.0038,
 45.8% identity in 48 aa overlap. Also similar to MTCY48_5
 and AL020958|SC4H8.7 Streptomyces coelicolor cosmid 4H8 (66
 aa), 41.0% identity in 39 aa overlap."
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein Rv0657c"
 /protein_id="CAB09377.1"
 /db_xref="GI:2143304"
 /db_xref="SPTREMBL:O06782"
 /translation="MSVTQIDLDELADVMRIAATVHTKKEAVNLAMRDYVERFRRIE
 ALARSR"
 complement(3691..3696)
 /note="possible RBS upstream of Rv0657c"
 complement(3761..4477)
 /gene="Rv0658c"
 complement(3761..4477)
 /gene="Rv0658c"
 /note="Rv0658c, (MTC1376.18), len: 238, unknown,

Job time : 709.59 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 : Search time 146.046 Seconds
(without alignments)
3438.621 Million cell updates/sec

Title: US-09-697-123B-13

Perfect score: 223

Sequence: 1 tcaaggagagcgctacgac.....ccgtgtgaaccgacgacat 223

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :				N_Geneseq_101002.*			
1:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*						
2:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*						
3:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*						
4:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*						
5:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*						
6:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*						
7:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*						
8:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*						
9:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*						
10:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*						
11:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*						
12:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*						
13:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*						
14:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*						
15:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*						
16:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*						
17:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*						
18:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*						
19:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*						
20:	/SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*						
21:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*						
22:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*						
23:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*						
24:	/SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*						

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	223	100.0	223	AAS05213	Mycobacterium gast
2	178.2	79.9	223	AAS05209	Mycobacterium scro
3	152.6	68.4	208	AAS05208	Mycobacterium kans
4	141.4	63.4	208	AAS05215	Mycobacterium aviu
5	138.2	62.0	208	AAS05217	Mycobacterium cels
6	135	60.5	208	AAS05201	Mycobacterium gord
7	135	60.5	208	AAS05219	Mycobacterium fort
8	135	60.5	208	AAS05224	Mycobacterium xeno
9	133.4	59.8	208	AAS05203	Mycobacterium gord

10	133.4	59.8	208	22	AAS05211	Mycobacterium mari
11	131.8	59.1	208	22	AAS05202	Mycobacterium gord
12	131.8	59.1	208	22	AAS05210	Mycobacterium ulce
13	126.6	56.8	207	22	AAS05204	Mycobacterium gord
14	123.8	55.5	208	22	AAS05205	Mycobacterium tube
15	123.8	55.5	208	22	AAS05206	Mycobacterium terr
16	123.8	55.5	208	22	AAS05216	Mycobacterium bovi
17	123.8	55.5	970	17	AAT09676	Mycobacterium tube
18	123.8	55.5	3519	22	AAH01976	Mycobacterium tube
19	123.8	55.5	3534	22	AAH02079	Mycobacterium tube
20	123.8	55.5	3853	21	AAA74651	M. tuberculosis rp
21	123.8	55.5	3853	21	AAH89994	Mycobacterium afri
22	122.2	54.8	208	22	AAS05222	Mycobacterium intr
23	121	54.3	205	22	AAS05220	Mycobacterium haem
24	120.6	54.1	208	22	AAS05223	Mycobacterium szul
25	118.2	53.0	207	22	AAS05212	M.leprae rpoB gene
26	111	49.8	3447	14	AAQ51532	Mycobacterium abs
27	107.8	48.3	214	22	AAS05221	Mycobacterium flav
28	104.4	46.8	211	22	AAS05218	Mycobacterium malm
29	101.4	45.5	214	22	AAS05214	Mycobacterium chel
30	99.8	44.8	214	22	AAS05207	C glutamicum codin
31	62.8	28.2	3495	22	AAH65512	C glutamicum codin
32	62.8	28.2	34980	22	AAH68525	Propionibacterium
33	52	23.3	27426	23	AAS59541	S. chrysomallus ac
34	42	18.8	3849	22	AAF25795	A. chrysogenum pho
35	41.2	18.5	3306	13	AAQ22202	Phosphoglycerate k
36	41.2	18.5	3306	13	AAQ23005	PGK. Acremonium c
37	41.2	18.5	3306	14	AAQ48534	Nucleotide sequenc
38	41.2	18.5	77536	21	AAA14651	Maize nitrite redu
39	41	18.4	278	24	ABL72018	Pseudomonas aerugi
40	39	17.5	471	23	AAS54080	Corr tassal-derive
41	39	17.5	5857	21	AAAS58471	Nucleotide sequenc
42	38.8	17.4	491	15	AAQ62119	Superoxide dismuta
43	37.8	17.0	1844	12	AAQ12000	Maize nitrite redu
44	37.2	16.7	1350	23	AAS54328	Pseudomonas aerugi
45	37.2	16.7	2760	16	AAQ79921	P. aeruginosa accB

ALIGNMENTS

RESULT 1
AAS05213
ID AAS05213 standard; DNA; 223 BP.

AC AAS05213:

DT 07-SEP-2001 (first entry)

DE Mycobacterium gastri rpoB gene fragment.

KW Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;

KW PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium gastri.

PN WO200131061-A1.

PD 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

PR 27-OCT-1999; 99KR-0046795.

PA (ERUM-) ERUME BIOTECH CO LTD.

PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

DR WPI: 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism


```
Db 1 TCAAGAGAAAGCGCTACGACCTGGCCCGCTGCGCCGCTACAAGGTCAACAAGAACTCG 60
QY 61 GCCTGAACACCGATCCCATCACCAACACGACGCTGACCGAAGAGAGGTGCTGCGCA 120
Db 61 GCCTGAACACCGAATGCGCCAAACACCAACGACCTGACCGAAGAGAGGTGCTGCGCA 120
QY 121 CCATCGAGTACCTGGTTCGCTCGACCAACGACCTCTCAGGGTGCCAGGCCCGCTTATGA 180
Db 121 CCATCGAATACCTGGTTCGCTCGACGAGG-----GGCAGCGCCACGATGA 165
QY 181 CTGTCCCGGGGGTGCAGGTGCGCGTGAACACCGACGACAT 223
Db 166 AGTCCCGGCTGCGCTGAGGTGCGCGTGGAGACCGACGACAT 208

RESULT 9
AAS05203
ID AAS05203 standard; DNA: 208 BP.
XX
AC AAS05203;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium gordonae type III rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium gordonae type III.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PK 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
DR WPI; 2001-300520/31.
XX
PS Claim 1; Page 41; 50pp; English.
XX
CC The present sequence for Mycobacterium gordonae type III rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 46 A; 69 C; 63 G; 30 T; 0 other;
```

```
Query Match 59.8%; Score 133.4; DB 22; Length 208;
Best Local Similarity 79.4%; Pred. No. 1.3e-22;
Matches 177; Conservative 0; Mismatches 31; Indels 15; Gaps 1;

QY 1 TCAAGAGAAAGCGCTACGACCTGGCCCGCTGCGCCGCTACAAGGTCAACAAGAACTCG 60
Db 1 TCAAGAGAAAGCGCTACGACCTGGCCCGCTGCGCCGCTACAAGGTCAACAAGAACTCG 60
QY 61 GCCTGAACACCGATCCCATCACCAACACGACGCTGACCGAAGAGAGGTGCTGCGCA 120
Db 61 GCCTGAACACCGATCCCATCACCAACACGACGCTGACCGAAGAGAGGTGCTGCGCA 120
QY 121 CCATCGAGTACCTGGTTCGCTCGACCAACGACCTCTCAGGGTGCCAGGCCCGCTTATGA 180
Db 121 CCATCGAGTACCTGGTTCGCTCGACGAGG-----CACGATGA 165
QY 181 CTGTCCCGGGGGTGCAGGTGCGCGTGAACACCGACGACAT 223
Db 166 CCGTCCGGGCGCACCGAGGTTCGCGTGGAGACCGACGACAT 208

RESULT 10
AAS05211
ID AAS05211 standard; DNA: 208 BP.
XX
AC AAS05211;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium marinum rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium marinum.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PK 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
DR WPI; 2001-300520/31.
XX
PS Claim 1; Page 43; 50pp; English.
XX
CC The present sequence for Mycobacterium marinum rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
```



```

DK WPI: 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism
XX
XX Disclosure: Page 41: 50pp: English.
XX
XX The present sequence for Mycobacterium tuberculosis rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
XX
Query Match 55.5%; Score 123.8; DB 22; Length 208;
Best Local Similarity 76.7%; Pred. No. 2.4e-20;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;

Qy 1 TCAAGGAGAAGCGCTACGACCTGGCCCGCGCTACAGGTCAACGAAGCTGG 60
Db 1 TCAAGGAGAAGCGCTACGACCTGGCCCGCGCTACAGGTCAACGAAGCTGG 60

Qy 61 CCTGAACACCGCATCATCCGATCACACACGACGCTGACCGAAGAAGACGTCGTCGCCA 120
Db 61 GCGTCGATGTCGGGAGGCCATCAGCTGCTGACGCTGACCGAAGAAGACGTCGTCGCCA 120

Qy 121 CCATCGAGTACCTGCTCGCTGCGACCGACGCTCTCAGGGTGGCCAGGCCCGCTTATGA 180
Db 121 CCATCGAATATCTGGTCCGCTTGCACGAGG-----GTCAGACCGACGATGA 165

Qy 181 CTGTCCCGGGGGTTCGAGGTGCGCGTGGAAACCGAGCAGCAT 223
Db 166 CCGTTCGGGCGGCTCGAGGTGCGCGTGGAAACCGAGCAGCAT 208

RESULT 15
AAS05206
ID AAS05206 standard; DNA; 208 BP.
XX
XX AAS05206;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium terrae rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria: rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
XX Mycobacterium terrae.
XX
XX WO200131061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX

```

```

PR 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ:
XX
XX WPI: 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism
XX
XX Claim 1: Page 42: 50pp: English.
XX
XX The present sequence for Mycobacterium terrae rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
XX
Query Match 55.5%; Score 123.8; DB 22; Length 208;
Best Local Similarity 76.7%; Pred. No. 2.4e-20;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;

Qy 1 TCAAGGAGAAGCGCTACGACCTGGCCCGCGCTACAGGTCAACGAAGCTGG 60
Db 1 TCAAGGAGAAGCGCTACGACCTGGCCCGCGCTACAGGTCAACGAAGCTGG 60

Qy 61 CCTGAACACCGCATCATCCGATCACACACGACGCTGACCGAAGAAGACGTCGTCGCCA 120
Db 61 GCGTCGATGTCGGGAGGCCATCAGCTGCTGACGCTGACCGAAGAAGACGTCGTCGCCA 120

Qy 121 CCATCGAGTACCTGCTCGCTGCGACCGACGCTCTCAGGGTGGCCAGGCCCGCTTATGA 180
Db 121 CCATCGAATATCTGGTCCGCTTGCACGAGG-----GTCAGACCGACGATGA 165

Qy 181 CTGTCCCGGGGGTTCGAGGTGCGCGTGGAAACCGAGCAGCAT 223
Db 166 CCGTTCGGGCGGCTCGAGGTGCGCGTGGAAACCGAGCAGCAT 208

Search completed: November 12, 2002, 16:50:12
Job time : 149.046 secs

```


QY 61 GCCTGAACACCGATCATCCGATCACCACACGACCGCTGACCGAAGAGAGCGTCTGTCGCCA 120
Db 86 GGTGATGTCGGCGAGCCCATCAGCTGTCGACGCTGACCGAAGAGAGCGTCTGTCGCCA 145
QY 121 CCATCGAGTACCTGGTTCGCTGACACACGCGCTCTCAGGTCGCCAGGCCCCCGTTATGA 180
Db 146 CCATCGAATATCTGCTCGCTTCGACGAGG-----GTCAGACACCGATGA 190
QY 181 CTGTCCCGCGGGTCTGAGTCCCGGTGCAACACCGACGACAT 223
Db 191 CCGTTCGGCGGGTCTGAGGTCGGGTGGAAACCGACGACAT 233

RESULT 2
PCT-US95-06790-1
: Sequence 1, Application PC/TUS9506790
: GENERAL INFORMATION:
: APPLICANT: Mayo Foundation for Medical Education and Research
: APPLICANT: and Hoffmann-La Roche Inc.
: TITLE OF INVENTION: Detection of a Genetic Locus Encoding
: TITLE OF INVENTION: Resistance to Rifampin
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Schwegman, Lundberg & Woessner
: STREET: 3500 IDS Center
: CITY: Minneapolis
: STATE: MN
: COUNTRY: USA
: ZIP: 55402

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/06790
: FILING DATE: 26-MAY-1995
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Raesch, Kevin W.
: REGISTRATION NUMBER: 35,651
: REFERENCE/DOCKET NUMBER: 150.105W01
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 612-339-0331
: TELEFAX: 612-339-3061
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 970 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA
: PCT-US95-06790-1

Query Match 55.5%; Score 123.8; DB 5; Length 970;
Best Local Similarity 76.7%; Pred. No. 1.4e-22;
Matches 171; Conservative 0; Mismatches 37; Indels 15; Gaps 1;
QY 1 TCAGGAGAACCGGTACGACCTGGCCGCGCTGCGCGCTACAAAGGTCAACAAAGAGCTGG 60
Db 26 TCAAGGAGAACCGGTACGACCTGGCCGCGCTGCGCTGCTAAGGTCAACAAAGAGCTGG 85
QY 61 GCCTGAACACCGATCATCCGATCACCACGACGCTGACCGAAGAGAGCGTCTGTCGCCA 120
Db 86 GGTGATGTCGGCGAGCCCATCAGCTGTCGACGCTGACCGAAGAGAGCGTCTGTCGCCA 145
QY 121 CCATCGAGTACCTGGTTCGCTGACACACGCGCTCTCAGGTCGCCAGGCCCCCGTTATGA 180
Db 146 CCATCGAATATCTGCTCGCTTCGACGAGG-----GTCAGACACCGATGA 190
QY 181 CTGTCCCGCGGGTCTGAGTCCCGGTGGAACCGACGACAT 223
Db 191 CCGTTCGGCGGGTCTGAGGTCGGGTGGAAACCGACGACAT 233

RESULT 3
US-08-313-185-57
: Sequence 57, Application US/08313185
: Patent No. 5851763
: GENERAL INFORMATION:
: APPLICANT: Heym, Beate
: APPLICANT: Cole, Stewart
: APPLICANT: Young, Douglas
: APPLICANT: Zhang, Ying
: APPLICANT: Honore, Nadine
: APPLICANT: Teienti, Amalio
: APPLICANT: Bodmer, Thomas
: TITLE OF INVENTION: Rapid Detection of Antibiotic Resistance
: TITLE OF INVENTION: in Mycobacterium Tuberculosis
: NUMBER OF SEQUENCES: 66
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSEE: Dunner
: STREET: 1300 I Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/313,185
: FILING DATE: 12-0CT-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Meyers, Kenneth J.
: REGISTRATION NUMBER: 25,146
: REFERENCE/DOCKET NUMBER: 02356.0068-00000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 408-4000
: TELEFAX: (202) 408-4400
: INFORMATION FOR SEQ ID NO: 57:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3447 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-313-185-57

Query Match 49.8%; Score 111; DB 2; Length 3447;
Best Local Similarity 73.1%; Pred. No. 2.5e-19;
Matches 163; Conservative 0; Mismatches 45; Indels 15; Gaps 1;
QY 1 TCAGGAGAACCGGTACGACCTGGCCGCGCTGCGCGCTACAAAGGTCAACAAAGAGCTGG 60
Db 809 TCAAGGAGAACCGGTACGACCTGGCCGCGCTTGTCTTACAAAGGTCAACAAAGAGCTGG 868
QY 61 GCCTGAACACCGATCATCCGATCACCACGACGCTGACCGAAGAGAGCGTCTGTCGCCA 120
Db 869 GGTGTCACGCGGTGAGTTGATCAGCTGCTCCACGCTGACCGAAGAGAGGATGTCTGTCGCCA 928
QY 121 CCATCGAGTACCTGGTTCGCTGCGCTGACCGACGCTCTCAGGTCGCCAGGCCCCCGTTATGA 180
Db 929 CCATAGAGTACCTGGTTCGCTGCTGCA-----TGAGGGTCACTCCGACATGA 973
QY 181 CTGTCCCGCGGGTCTGAGGTCGGGTGGAAACCGACGACAT 223
Db 974 CTGTCCAGGTGGGTAGAAAGTCCAGTGCAGAACTGACGATAT 1016

RESULT 4
US-09-082-614A-57

QY	61	GCTTGAAACCCGATCATCGATCACCACCAAGACGCCTCAACCGAAGAAGACGTCCTGCCCA	120
Db	1515	GCATCTGTGGTGGCTTCCCACCATCAAGTAGCGCGTCGAGCATGGGCCAAGGCCGTCA	1574
QY	121	CCATCCAGTACTGGTTTCGGCTGCACCAAGCCTCTCAGGTGGCGCAGGCCGCCCTTATGA	180
Db	1575	TCTCATGTCCCCACCTTGGCGCGCCCAAGCGGACCCCAACCCCAAGTACTCGCTGCAGC	1634
QY	181	CTGTCCCGCGGGGTCGAGGTGCGCGTGGAAACCGACGACA	222
Db	1635	CCGTCGTCCCGCAGCTCGAGAAGCTGCTCGGCAAGAGGGTCA	1676

RESULT 6

```

US-08-133-711-36
: Sequence 36, Application US/08133711
: Patent No. 5252463
: GENERAL INFORMATION:
: APPLICANT: Zolig, Werner
: TITLE OF INVENTION: Methods and reagents for detecting
: TITLE OF INVENTION: pathogens using superoxide d
: TITLE OF INVENTION: targeting
: NUMBER OF SEQUENCES: 44
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hoffmann-La Roche Inc.
: STREET: 340 Kingsland Street
: CITY: Nutley
: STATE: NJ
: COUNTRY: U.S.A.
: ZIP: 07110-1199
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION NUMBER: US/08/133,711
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: EP 92810780.4
: FILING DATE: 13-OCT-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Patricia S. Rocha
: REGISTRATION NUMBER: 31,054
: REFERENCE/DOCKET NUMBER: 4095/95
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (201) 325 5000
: TELEFAX: (201) 235 3500
: INFORMATION FOR SEQ ID NO: 36:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 491 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: ORIGINAL SOURCE:
: ORGANISM: Mycobacterium kansasii/SOD gene
US-08-133-711-36

```

```

Query Match      17.4%; Score 38.8; DB 1; Length 491;
Best Local Similarity 48.6%; Pred. NO. 0.19;
Matches 106; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```

Oy	1	TCAAGGAGAACGGCTTACGACCTTGGCCCGCTTCCGCGCTTACAAGGTCAACAAGAGCTGG	60
Db	100	TCATTGCTGAACGAGAGAACTTTGGCTTCAACCTTCGCGCGCCACGTCAACACACACGATCT	159
Oy	61	GCCTGAACACCGATCATCCGATCAACACACGACGCTGACCGAGAGAGACGTGCTGCCTCA	120
Db	160	GGTGAAGAAGACTTTCTCCCAACGAGGCGCACAAAGCGACCGCGCAACTCGCGCGCGCCA	219

Qy	121	CCATCGAGTACCTGGTTTCGGCTCGACACGCCCTCTCAGGTGGCCAGGCCGCCCGTTATGA	180
Db	220	TCGACGAGGGGTTTCGGGTCTCTCGACAAGTTTCGTGCCAAATTCACGCCGCCGCCACCA	279
Qy	181	CTGTCCCGGGGGGTTCAGGTCCCGTGGAAACCGAC	218
Db	280	CGGTCCAGGGTTCGGGTTCGGCGCGCTGGGCTGGGAC	317

RESULT 7

```

US-08-074-121-4
: Sequence 4, Application US/08074121
: Patent NO. 5767362
: GENERAL INFORMATION:
: APPLICANT: Best, Elaine
: APPLICANT: Knauf, Vic C.
: TITLE OF INVENTION: Methods and Compositions for Modulating
: TITLE OF INVENTION: Lipid Content of Plant Tissues
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson
: STREET: 2200 Sand Hill Road, Suite 100
: CITY: Menlo Park
: STATE: California
: COUNTRY: US
: ZIP: 94025
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/074,121
: FILING DATE: 08-JUN-1993
: CLASSIFICATION: 800
: ATTORNEY/AGENT INFORMATION:
: NAME: Rae-Venter, Barbara
: REGISTRATION NUMBER: 32,750
: REFERENCE/DOCKET NUMBER: 05938/043001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 854-5277
: TELEFAX: (415) 854-0875
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3231 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 861..1328
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1349..2695
: US-08-074-121-4

```

Query Match	16.7%	Score 37.2;	DB 1;	Length 3231;
Best Local Similarity	51.2%	Pred. No. 0.61;		
Matches 87; Conservative	0;	Mismatches 83;	Indels 0;	Caps 0;

[illegible]

```

RESULT 8
PCT-US94-06447-4
: Sequence 4, Application PC/TUS9406447
: GENERAL INFORMATION:
: APPLICANT: Calgene, Inc.
: TITLE OF INVENTION: Methods and Compositions for Modulating
: TITLE OF INVENTION: Lipid Content of Plant Tissues
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESS: Weill, Gotschal & Mangos
: STREET: 2882 Sand Hill Road, Suite 280
: CITY: Menlo Park
: STATE: California
: COUNTRY: US
: ZIP: 94025
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/06447
: FILING DATE: 06-JUN-1994
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Rae-Venier, Barbara
: REGISTRATION NUMBER: 32,750
: REFERENCE/DOCKET NUMBER: CGNE-097/WO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 926-6200
: TELEFAX: (415) 854-3713
: TELEX:
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3231 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 861..1328
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1349..2695
PCT-US94-06447-4

```

	Query Match	16.7%	Score 37.2:	DB 5:	length 3231:
	Best Local Similarity	51.2%:	Pred. No. 0.61:		
	Matches 87:	Conservative	0:	Mismatches 83:	Indels 0:
	Gaps	0:			
QY	29	CGTCGGCGCTACAAAGCTCAACAAGAAGCTCGGCGCTCAACACCCGATCATTCGATCACCAC	88		
Db	1996	CGACGGCCAGGCAACGCCATCCACCTCGCGGACCGGGAGCTGCTCCCTGCAGCGCCGCCA	2055		
QY	89	CACGACGCTCACCAAGAAGACGTCGTCGCCACCATCGAGTACCTGGTTCGCCCTGCACCA	148		
Db	2056	CCAGAAGGTGATCGAAGAGCGCCCGGCCGCGATCGACGAGAAGCTGCCCAGGAAGT	2115		
QY	149	GGCTCTCAGGTGGCCAGGCCCGCTTATGACTGTCCCGCGGGGTGCG	198		
Db	2116	CTTTCGCGCGTGCCTCCAGCGTCCATTCGATCGGTACGCGCGCGCG	2165		

RESULT 9
US-09-221-017B-1070
; Sequence 1070, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: ROSS, BRUCE C.
; TITLE OF INVENTION: P. GINGIVALIS N

```

1 NUMBER OF SEQUENCES: 1120
2 CORRESPONDENCE ADDRESS:
3 ADDRESSEE: MORRISON & FOERSTER
4 STREET: 755 PAGE MILL ROAD
5 CITY: Palo Alto
6 STATE: CA
7 COUNTRY: USA
8 ZIP: 94304-1018
9 COMPUTER READABLE FORM:
10 MEDIUM TYPE: Diskette
11 COMPUTER: IBM Compatible
12 OPERATING SYSTEM: Windows
13 SOFTWARE: FastSO for Windows Version 2.0b
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/09/221,017B
16 FILING DATE: 23-DEC-1998
17 CLASSIFICATION:
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: PP1182
20 FILING DATE: 31-DEC-1997
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: PP1546
23 FILING DATE: 30-JAN-1998
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: PP2911
26 FILING DATE: 09-APR-1998
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: PCT/AU98/01023
29 FILING DATE: 10-DEC-1998
30 ATTORNEY/AGENT INFORMATION:
31 NAME: Monroy, Gladys H
32 REGISTRATION NUMBER: 32,430
33 REFERENCE/DOCKET NUMBER: 27340-20021.00
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE: 650-813-5600
36 TELEFAX: 650-494-0792
37 TELEX: 706141
38 INFORMATION FOR SEQ ID NO: 1070:
39 SEQUENCE CHARACTERISTICS:
40 LENGTH: 734 base pairs
41 TYPE: nucleic acid
42 STRANDEDNESS: double
43 TOPOLOGY: circular
44 MOLECULE TYPE: DNA (genomic)
45 HYPOTHETICAL: NO
46 ANTI-SENSE: UNKNOWN
47 ORIGINAL SOURCE:
48 ORGANISM: PORYPHYROMONAS GINGIVALIS
49 FEATURE:
50 NAME/KEY: misc_feature
51 LOCATION: 1..734
52 US-09-221-017B-1070

```

	Query Match	16.4%	Score 36.6:	DB 4:	Length 734:	
	Best Local Similarity	59.9%:	Pred. No. 0.71:	Mismatches 0:	Indels 49:	Gaps 6:
	Matches 82:	Conservative				
Oy	1	TCAAGGAGAAGCGGTACGACCCTGGCCCCGCTCGGCCGCTACAAGTCCAAGAAAGCTGG	60			
Db	4	TCTCTGAATAAACGATACGACTTTGGCCGATTCGGACGCCTATCGTATCATTAATAAGAAACTCA	63			
Oy	61	GCCTTGACACCGCATCCGATCACCCACGAGCGCTCACCAGAAGAACGCTCGTCGGCCA	120			
Db	64	ATTTCGATATTGATCCCTGATATCAAAG-----TGCTGACCAANTGAAGATATTATTGAGA	117			
Oy	121	CCATCGAGTAGCTGGTT	137			
Db	118	TCATCAAGTATCTTATT	134			

RESULT 10
US-08-718-904-79
; Sequence 79, Application US/08718904

```

: Patent No. 6037329
:
: GENERAL INFORMATION:
:
: APPLICANT: Baird, J. Andrew
: APPLICANT: Chandler, Lois Ann
: APPLICANT: Sosnowski, Barbara A.
: TITLE OF INVENTION: COMPOSITIONS CONTAINING NUCLEIC ACIDS AND LIGANDS FOR THERAPE
:
: NUMBER OF SEQUENCES: 128
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: SEED and BERRY LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98104-7092
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/08/718,904
: FILING DATE: 24-SEP-1996
: CLASSIFICATION: 424
:
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 6037329tenburg Ph.D., Carol
: REGISTRATION NUMBER: 39,317
: REFERENCE/DOCKET NUMBER: 760100.415C1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
:
: INFORMATION FOR SEQ ID NO: 79:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 765 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: FEATURE:
:
: NAME/KEY: CDS
: LOCATION: 1..762
: OTHER INFORMATION: /product= "Mammalian codon
: OTHER INFORMATION: optimized saporin"
:
: US-08-718-904-79

```

	Query Match	16.4%;	Score 36.6;	DB 3;	Length 765;
	Best Local Similarity	53.1%;	Pred. No. 0.71;		
	Matches	78; Conservative	0; Mismatches	69; Indels	0; Gaps
Qy	6	GAGAAAGCGCTACGACCTGGCCGCCGCTGCAGAGTCAACAAGGTCAACGAAGAGCTGGGCGCTG	65		
Db	208	GACAACCTGTACGTGGTGCGCTACCTGGCCATGAGCAACCAACGTAACCGCGCCTAC	267		
Qy	66	AACACCAGTCATCCGATCACCAACGACGACGCTGACCGAAGAAGACGTCGTGGCCACCATC	125		
Db	268	TACTTCAAGTCCGNATACACCTCGCGCGAGCTACCGCCCTGTTCCTTAGGGCCACCAC	327		
Qy	126	GAGTACCTGGTTGCCGTGCACCAAGCC	152		
Db	328	GCCAACCAAGAGGCCCTTGGAGTACACC	354		

```

RESULT 11
US-08-972-258-1
: Sequence 1, Application US/08972258
: Patent No. 5929301
: GENERAL INFORMATION:
: APPLICANT: Baszczynski, Chris
: APPLICANT: Bowen, Benjamin A.
: APPLICANT: Drummond, Bruce J.
: APPLICANT: Gordon-Kamm, William J.
: APPLICANT: Peterson, David J.
: APPLICANT: Sandahl, Gary A.
: APPLICANT: Tagliani, Laura A.
: APPLICANT: Zhao, Zuo-Yu

```

```

; TITLE OF INVENTION: No. 5929301el Nucleic Acid Sequence Encoding FLIP
; Patent No. 5929301
; TITLE OF INVENTION: Recombinase and Method of Using Same
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: US
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/972,258
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2202
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1272 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Synthetic sequence (optimized)
; US-08-972-258-1

```

Query Match 16.2%; Score 36.2; DB 2; Length 1272;
Best Local Similarity 48.8%; Pred. No. 0.96;
Matches 98; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy	20	CCTGGCCGGGTGGCCGCTTACAAGGTCAACAGAGACTGGGCTGAACACCCGATCATCC	79
Db	540	CCTTCTCTGCCACCTTCATCAACTGCGCAGGTTCTCAGACATCAAGAAACGTGGACCC	599
Qy	80	GATCACACACGAGCGTGACCCGAAGAAGAGCTGTCGCCACCATCGAGTACCTGCTTCG	139
Db	600	CAAGTCCTTCAAGTCTCGTGAGAACAGTACCTCGGCGTGATCATCCAGTGCCTCGTGAC	659
Qy	140	CCTGCACACGCGCTTCAGGGTGGCCAGGCCCGCCGTTATGACTGTCCCGGGGGGTGCA	199
Db	660	CGAGACCAAGACCTCCGTGTCAGGCACATCTACTTCTCTCCGCTCGCGCAGGATCGA	719
Qy	200	GGTCCCGGTGGAACACGACGA	220
Db	720	CCCCCTCGTGTACTCTCGACGA	740

```

RESULT 12
US-09-263-128-1
; Sequence 1, Application US/09263128
; Patent No. 6175058
; GENERAL INFORMATION:
; APPLICANT: Baszczynski, Chris
; APPLICANT: Bowen, Benjamin A.
; APPLICANT: Drummond, Bruce J.
; APPLICANT: Gordon-Kamm, William
; APPLICANT: Peterson, David J.
; APPLICANT: Sandahl, Gary A.
; APPLICANT: Tagliani, Laura A.
; APPLICANT: Zhao, Zuo-Yu
; TITLE OF INVENTION: NO 6175058

```


OTHER INFORMATION: Description of Artificial Sequence: sequence
OTHER INFORMATION: encoding moCre:FLPm, Cre from Bacteriophage P1 and
OTHER INFORMATION: FLP from Saccharomyces, both maize preferred
OTHER INFORMATION: codons
NAME/KEY: CDS
LOCATION: (1)..(2346)
US-09-193-503B-5

Query Match 16.2%; Score 36.2; DB 4; Length 2346;
Best Local Similarity 48.8%; Prod. No. 1;
Matches 98; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
QY 20 CCTGCGCCGGTCGGCGCTACAAAGGTCAACAAAGAGCTGGGCTGAACACCGATCATCC 79
DB 1614 CCCTCTCTCCGCCACCTTCATCAACTCCGGCAGGTTCTCAGACATCAAGAACCTGGACCC 1673
QY 80 GATCACCACACGACGCTGACCAAGAGAGAGCTGCTGGCCACCATCGAGTACCTGGTTGG 139
DB 1674 CAAGTCCCTTCAAGCTGCTGAGAACAAAGTACCTCGGCGGTGATCATCGAGTGCCTCGTGAC 1733
QY 140 CCTGCACACGCGCTCTCAGGCTGCGCAGGCCCGCGTTTATGACTGTCCCGCGGGGTGCGA 199
DB 1734 CGACACCAAGACCTCCGTGTCAGGCACATCTACTTCTTCTCCGCTCGCGCGAGGATCGA 1793
QY 200 GGTCCGGTGGAAACCGACGA 220
DB 1794 CCCCCTCGTGTACCTCGACGA 740

RESULT 15

US-09-193-503B-8
Sequence 8, Application US/09193503B
Patent No. 6262341
GENERAL INFORMATION:
APPLICANT: Baszczynski, Christopher L.
APPLICANT: Lyznik, Leszek A.
APPLICANT: Gordon-Kamm, William J.
APPLICANT: Guan, Xueni
APPLICANT: Rao, Guru
APPLICANT: Tagliani, Laura A.
TITLE OF INVENTION: A No. 6262341el Method For The Integration Of Foreign DNA Into
TITLE OF INVENTION: Eukaryotic Genomes
FILE REFERENCE: 5718-66 (amended listing)
CURRENT APPLICATION NUMBER: US/09/193,503B
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/099,435
PRIOR FILING DATE: 1998-09-08
PRIOR APPLICATION NUMBER: 60/056,627
PRIOR FILING DATE: 1997-11-18
PRIOR APPLICATION NUMBER: 60/065,613
PRIOR FILING DATE: 1997-11-18
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 2346
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Sequence
OTHER INFORMATION: encoding a FLPm:Cre polypeptide, FLP from
OTHER INFORMATION: Saccharomyces (maize preferred codons), and Cre
OTHER INFORMATION: from Bacteriophage P1
NAME/KEY: CDS
LOCATION: (1)..(2346)
US-09-193-503B-8

Query Match 16.2%; Score 36.2; DB 4; Length 2346;
Best Local Similarity 48.8%; Prod. No. 1;
Matches 98; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 20 CCTGCGCCGGTCGGCGCTACAAAGCTCAACAAAGAGCTGGGCTGAACACCGATCATCC 79
DB 540 CCTCTCTCCGCCACCTTCATCAACTCCGGCAGGTTCTCAGACATCAAGAACGTGGACCC 599

QY 80 GATCACCACGAGCGCTGACCCGAAGACAGACGTCGTCCGCCACCATCGAGTACCTGGTTGG 139
DB 600 CAAGTCCTTCAAGCTCGTGCAGAACAAAGTACCTCGGCGGTGATCATCCAGTGCCTCGTGAC 659
QY 140 CCTGCACACGCGCTCTCAGGCTGAGGGTGGCCAGGCCCGCGTTTATGACTGTCCCGCGGGGGGTGCA 199
DB 660 CGAGACCAAGAGCTCGGTGTCCAGGCACATCTACTTCTTCTCCGCTCGCGCGGAGGATCGA 719
QY 200 GGTCCGGTGGAAACCGACGA 220
DB 720 CCCCCTCGTGTACCTCGACGA 740

Search completed: November 12, 2002, 20:14:30
Job time : 51.9052 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 16:05:40 ; Search time 22.6867 Seconds
(without alignments)
3487.380 Million cell updates/sec

Title: US-09-697-123B-13

Perfect score: 223

Sequence: 1 tcaagagagactacgac.....coggtgaaacacgacat 223

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DH seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications, NA: *

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	62.8	28.2	5096	10	US-09-984-711-5	Sequence 5, Appli
2	62.8	28.2	5099	9	US-10-075-460-5	Sequence 5, Appli
3	62.8	28.2	5099	10	US-09-887-052-1	Sequence 1, Appli
4	62.8	28.2	5099	10	US-09-887-052-3	Sequence 3, Appli
5	62.8	28.2	5099	10	US-09-887-052-5	Sequence 5, Appli
6	42.4	19.0	704	12	US-10-062-254-109	Sequence 109, App
7	42.4	19.0	752	12	US-10-062-254-111	Sequence 111, App
8	41	18.4	278	10	US-09-294-093B-1392	Sequence 1392, Ap
9	39	17.5	471	10	US-09-815-242-7717	Sequence 7717, Ap
10	38	17.0	175	10	US-09-923-876-5702	Sequence 5702, Ap
11	37.4	16.8	626	12	US-10-062-254-115	Sequence 115, App
12	37.2	16.7	1350	10	US-09-815-242-7965	Sequence 7965, Ap
13	36.4	16.3	1806	10	US-09-815-242-4015	Sequence 4015, Ap
14	36	16.1	642	12	US-10-062-254-123	Sequence 123, App
15	36	16.1	918	10	US-09-815-242-7705	Sequence 7705, Ap
16	35.8	16.1	1535	12	US-10-052-586-545	Sequence 545, App
17	35.6	16.0	969	10	US-09-815-242-7933	Sequence 7933, Ap
18	34.8	15.6	1206	10	US-09-815-242-7692	Sequence 7692, Ap
19	34.4	15.4	529	12	US-10-062-254-107	Sequence 107, App

ALIGNMENTS

RESULT 1

US-09-984-711-5
; Sequence 5, Application US/09984711
; Patent No. US20020119549A1

; GENERAL INFORMATION:
; APPLICANT: MOECKEL, Bettina

; APPLICANT: BATHE, Brigitte

; APPLICANT: STEPHAN, Hans

; APPLICANT: KREUTZER, Caroline

; APPLICANT: HERMANN, Thomas

; APPLICANT: PFEFFERLE, Walter

; APPLICANT: BINDER, Michael

; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpsL GENE

; FILE REFERENCE: 204209USO

; CURRENT APPLICATION NUMBER: US/09/984,711

; CURRENT FILING DATE: 2001-10-31

; PRIOR APPLICATION NUMBER: DE10108230.9

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 5

; LENGTH: 5096

; TYPE: DNA

; ORGANISM: Corynebacterium glutamicum

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (702)..(4196)

; OTHER INFORMATION:

US-09-984-711-5

Query Match 28.2% Score 62.8; DB 10; Length 5096;

Best Local Similarity 71.0%; Pred.No. 6.1e-08;

Matches 98; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

Qy 9 AAGCGCTACGACCTGGCCGCGCTGCGGCTTACAAGGTCAACAAGAGCTGGGCTGAAC 68

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 1578 AAGCGCTACGACCTGGCTGCGGTTGCGTTACAAAGATCAACCGCAAGCTCGGCTTGGT 1637

Qy 69 ACCGATCATCGGTCACCAACGACGCTGACCAAGAGAGCTGCGCGCCACCATCGAG 128

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 1638 GCGCACCACGATGGTTTGTAT---GACTCTTACTGAAGGAGACATCGCAACCATCGAG 1694

GENERAL INFORMATION:
APPLICANT: MOECKEL, Bettina
APPLICANT: BATHE, Brigitte
APPLICANT: HERMANN, Thomas
APPLICANT: PFEFFERLE, Walter
APPLICANT: BINDER, Michael
TITLE OF INVENTION: NUCLEOTIDE SEQUENCES WHICH CODE FOR THE rpoB GENE:
FILE REFERENCE: 204212USOX
CURRENT APPLICATION NUMBER: US/09/887.052
CURRENT FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: DE10107229.5
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 5099
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: CDS
LOCATION: (702)..(4196)
US-09-887-052-5

Query Match 28.2%; Score 62.8; DB 10; Length 5099;
Best Local Similarity 71.0%; Pred. No. 6.1e-08;
Matches 98; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

QY 9 AAGCGCTACGACCTGGCCGCTGCGCTACAGGTCAACAAGAGTCAACGAGCTGGGCTGAAC 68
|||||
Db 1578 AAGCGCTACGACCTGGCTGCTTACAGATCAACCGCAAGCTCGGCTTGGT 1637
|||||

QY 64 ACCGATATCCGATCACCACACGACGCTGACCGAAGAGAGCTGTCGCCACCATCGAG 128
|||||
Db 1638 GCGCACCAAGATGGTTTTCAT---GACTCTTACTGAAGAGGACATCGCAACCAACCATCGAG 1694
|||||

QY 129 TACCTGGTTCGCTGCAC 146
|||||
Db 1695 TACCTGGTTCGCTGCAC 1712
|||||

RESULT 6
US-10-062-254-109
Sequence 109, Application US/10062254
Patent No. US20020138882A1
GENERAL INFORMATION:
APPLICANT: Cahoon, Edgar B
APPLICANT: Cahoon, Rebecca E
APPLICANT: Falco, Saverio Carl
APPLICANT: Fang, Yiwen
APPLICANT: Hantke, Sabine S.
APPLICANT: Lee, Jian-Ming
APPLICANT: Li, Zhongsen
APPLICANT: Miao, Guo-Hua
APPLICANT: Morgante, Michele
APPLICANT: Niu, Xiping
APPLICANT: Odell, Joan
APPLICANT: Rafalski, Antoni
APPLICANT: Sakai, Hajime
APPLICANT: Zheng, Peizhong
APPLICANT: Zhu, Qun
TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/062,254
CURRENT FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 09/630,346
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 60/146511
PRIOR FILING DATE: 1999-07-30
PRIOR APPLICATION NUMBER: 60/156006
PRIOR FILING DATE: 1999-09-23
PRIOR APPLICATION NUMBER: 60/156899
PRIOR FILING DATE: 1999-09-30
PRIOR APPLICATION NUMBER: 60/157287

PRIOR FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: 60/169767
PRIOR FILING DATE: 1999-12-09
PRIOR APPLICATION NUMBER: 60/171054
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: 60/172958
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/171515
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: 60/173535
PRIOR FILING DATE: 1999-12-29
NUMBER OF SEQ ID NOS: 375
SOFTWARE: Microsoft Office 97
SEQ ID NO 109
LENGTH: 704
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (640)
NAME/KEY: unsure
LOCATION: (663)
NAME/KEY: unsure
LOCATION: (688)
NAME/KEY: unsure
LOCATION: (692)
NAME/KEY: unsure
LOCATION: (696)
US-10-062-254-109

Query Match 19.0%; Score 42.4; DB 12; Length 704;
Best Local Similarity 58.9%; Pred. No. 0.0094;
Matches 73; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 4 AGGAGAGCGCTAGCAGCTGGCCGCTGCGCGCTCAAGGTCAACAAGAGCTGGGCC 63
|||||
Db 275 AGCTGATGCGCGCAGCAGCTGCCACCGATGCCGCGCAGGAGGTCCCAAGAGGTGGGCC 334
|||||

QY 64 TGAACACCGATCATCCGATCACCACGACGCTGACCGAAGAGAGCTGTCGCCACCA 123
|||||
Db 335 TCGTCTACCGAGCTCAACATCGCCCGCCCAAGAGATCGGCGTTGATGAGGAGATCTTCGTGG 394
|||||

QY 124 TCGA 127
|||||
Db 395 TCGA 398
|||||

RESULT 7
US-10-062-254-111
Sequence 111, Application US/10062254
Patent No. US20020138882A1
GENERAL INFORMATION:
APPLICANT: Cahoon, Edgar B
APPLICANT: Cahoon, Rebecca E
APPLICANT: Falco, Saverio Carl
APPLICANT: Fang, Yiwen
APPLICANT: Hantke, Sabine S.
APPLICANT: Lee, Jian-Ming
APPLICANT: Li, Zhongsen
APPLICANT: Miao, Guo-Hua
APPLICANT: Morgante, Michele
APPLICANT: Niu, Xiping
APPLICANT: Odell, Joan
APPLICANT: Rafalski, Antoni
APPLICANT: Sakai, Hajime
APPLICANT: Zheng, Peizhong
APPLICANT: Zhu, Qun
TITLE OF INVENTION: Polynucleotides Encoding Proteins Involved In Plant Metabolism
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/062,254
CURRENT FILING DATE: 2002-02-01
PRIOR APPLICATION NUMBER: 09/630,346
PRIOR FILING DATE: 2000-07-28

; PRIOR APPLICATION NUMBER: 60/146511
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/156006
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/156899
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 60/157287
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 60/169767
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 60/171054
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/172958
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/171515
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: 60/171535
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 375
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 111
; LENGTH: 752
; TYPE: DNA
; ORGANISM: Zea mays
; US-10-062-254-111

Query Match
Best Local Similarity 19.0%; Score 42.4; DB 12; Length 752;
Matches 73; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 4 AGGAGAAGCGCTACGACCTGGCCCGCTGCGCGCTTACAAAGTCAACAAAGAGCTGGGCC 63
Db 236 AGCTGATGCGCCAGCAGCTGGCCAGCGATGCGCGCAAGGAGGTCAACAAAGAGGTGGGCC 295
QY 64 TGAACACCGCATCCGATCACCACCGCTGACCGTGAACGAGAGCTCGTCGCCACCA 123
Db 296 TCGTCTACCACTCAACATCGCCCAAGAGAGTCTGATGAGGAGATCTTCGTGG 355
QY 124 TCGA 127
Db 356 TCGA 359

RESULT 8
US-09-294-093B-1392
; Sequence 1392, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath, V.
; APPLICANT: Ito, Iaura, Y.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
; FILE REFERENCE: PL-0009 US
; CURRENT APPLICATION NUMBER: US/09/294,093B
; PRIOR FILING DATE: 1999-04-16
; PRIOR APPLICATION NUMBER: 60/082,567
; PRIOR FILING DATE: April 21, 1998
; NUMBER OF SEQ ID NOS: 6207
; SOFTWARE: PERL Program
; SEQ ID NO 1392
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20010051335A1 700344235H1
; NAME/KEY: unsure
; LOCATION: 158
; OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-1392

Query Match
Best Local Similarity 18.4%; Score 41; DB 10; Length 278;
Matches 73; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

Matches 68; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 4 AGGAGAAGCGCTACGACCTGGCCCGCTGCGCGCTTACAAAGTCAACAAAGAGCTGGGCC 63
Db 164 AGCTGATGCGCCAGCAGCTGGCCAGCGATGCGCGCAAGGAGGTCAACAAAGAGGTGGGCC 223
QY 64 TGAACACCGCATCCGATCACCACCGCTGACCGTGAACGAGAGCTCGTC 116
Db 224 TCGTCTACCACTCAACATCGCCCAAGAGATCGCGGTTGATGAGGAGATC 276

RESULT 9
US-09-815-242-7717
; Sequence 7717, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7717
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(471)
US-09-815-242-7717

Query Match
Best Local Similarity 17.5%; Score 39; DB 10; Length 471;
Matches 102; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 1 TCAAGGAGAAGCGCTACGACCTGGCCCGCTGCGCGCTTACAAAGTCAACAAAGAGCTGG 60
Db 86 TCTACCTGAAAACCCACAACCTTCCACTTGAAGCTCACCGCGCCGATGTTCAACAGCTTGC 145
QY 61 GCCTGAACACCGCATCCGATCACCACCGCTGACCGTGAACGAGAGCTCGTCGCCCA 120
Db 146 ACTTGATGTTGAAAGGCCAGTACACCGAACTGGCCGTTGCCGTGGACGACATCGCCGAGC 205
QY 121 CCATCGAGTACCTGGTTTCGCGCTGCACCGCCTCTCAGGGTGGCCAGGCCCCCGTTATGA 180
Db 206 GCATCCGGCCCTGGGCTTCCCGGCGCGGACCTACGCCGCCCTAGCGCCGCTGCTCT 265
QY 181 CTGTCCCGCGGGGTTCGAGGTGCCGG 207
Db 266 CCATCAAGGAGGAGGAAGCGGTTCCGG 292


```
: LENGTH: 642
: TYPE: DNA
: ORGANISM: Trilicium aestivum
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (34)
: NAME/KEY: unsure
: LOCATION: (624)
US-10-062-254-123

Query Match          16.1%   Score 36;   DB 12;   Length 642;
Best Local Similarity 55.6%   Pred. No. 0.43;
Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy  4 AGGAGAAGCGGTACGACCTGGCCCGCGTGGCCGCTACAAAGGTCAACAAGAGCTGGGCC 63
    ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db  236 AGCAGATGAGCGCCACCTCGCCACCGACCGCGCGGCGCAAGGAGATCGCCAAAGGTCGGCT 295

Qy  64 TGAACACCGATATCCGATCACCATCCAGCGCTGACCGAAGAGAGCTGCTCCGCCACCA 123
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  296 TCGTCTACCAAGTCAACATCTCGCCCAAGAAGATGGGCGTGAGGAGATCTTCGTGCG 355

Qy  124 TCGA 127
    | |
Db  356 TCGA 359

RESULT 15
US-09-815-242-7705
: Sequence 7705, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Kari L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA-011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7705
: LENGTH: 918
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(918)
US-09-815-242-7705

Query Match          16.1%   Score 36;   DB 10;   Length 918;
Best Local Similarity 47.7%   Pred. No. 0.45;
Matches 105; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
```

```
Qy  2 CAAGGAGAAGCGGTACGACCTGGCCCGCGTGGCCGCTACAAAGGTCAACAAGAGCTGGG 61
    ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Db  180 CGAGGTGCAGGCGGTCTACGTTCGATACGCCCGGCTCGCAAGAGCGCGGAGAGGCCCT 239

Qy  62 CCTGAACACCGATCATCCGATCACCACCACGAGCGCTGACCGAAGAGAGCGTCTGTCGCCAC 121
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  240 CAACCGCTACATGAACCGTACCGCCACGCGGCGCATTTGAAGGATGTCGACGCTGTGATCTT 299

Qy  122 CATCGAGTACCTGGTTTCGCTTGACACACGCGCTCTCAGGGTGGCCAGGCCGCCGTTATGAC 181
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  300 CGTGGTCGACCGCACCCCGCTGGACCGAAGAGGACCGACATGCTCTGGAGCGCGTGCAGTA 359

Qy  182 TGTCCCCCGCGGGTTCGAGGTGCCGCTGGAACCGGACCGAC 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  360 CGTCAGTTGCCCGGTGCTGATCGCGGTGAACAAGACCGAC 399
```

Search completed: November 12, 2002, 16:58:16
Job time : 33.6867 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 12, 2002, 15:44:35 : Search time 136.222 Seconds
(without alignments)
3438.621 Million cell updates/sec

Title: US-09-697-123B-8
Perfect score: 208
Sequence: 1 tcaagagaagcgctacgac.....ccggtgaaaccgacgacat 208

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	ID	Description	
1	208	100.0	208	22	AAS05208	Mycobacterium kans
2	174.4	83.8	208	22	AAS05217	Mycobacterium celi
3	171.2	82.3	208	22	AAS05215	Mycobacterium aviu
4	169.6	81.5	208	22	AAS05203	Mycobacterium gord
5	169.6	81.5	208	22	AAS05211	Mycobacterium mari
6	168	80.8	208	22	AAS05201	Mycobacterium gord
7	168	80.8	208	22	AAS05219	Mycobacterium fort
8	166.4	80.0	208	22	AAS05205	Mycobacterium tube
9	166.4	80.0	208	22	AAS05206	Mycobacterium terr

10	166.4	80.0	208	22	AAS05210	Mycobacterium ulce
11	166.4	80.0	208	22	AAS05216	Mycobacterium bovi
12	166.4	80.0	970	17	AAT09676	Mycobacterium tube
13	166.4	80.0	3519	22	AAH51976	Mycobacterium tube
14	166.4	80.0	3534	22	AAH02079	Mycobacterium tube
15	166.4	80.0	3853	21	AAA74651	M. tuberculosis rp
16	166.4	80.0	3853	21	AAA89994	M. tuberculosis rp
17	164.8	79.2	208	22	AAS05222	Mycobacterium afri
18	161.6	77.7	208	22	AAS05224	Mycobacterium xeno
19	160	76.9	208	22	AAS05202	Mycobacterium gord
20	159.2	76.5	207	22	AAS05204	Mycobacterium gord
21	158.4	76.2	208	22	AAS05223	Mycobacterium haem
22	152.6	73.4	223	22	AAS05213	Mycobacterium gast
23	151.2	72.7	207	22	AAS05212	Mycobacterium szul
24	150.4	72.3	205	22	AAS05220	Mycobacterium intr
25	142.4	68.5	214	22	AAS05221	Mycobacterium absc
26	142.4	68.5	3447	14	AAQ51532	M.leprae rpoB gene
27	139.2	66.9	214	22	AAS05218	Mycobacterium maln
28	134.2	64.5	211	22	AAS05214	Mycobacterium flav
29	132.8	63.8	214	22	AAS05207	Mycobacterium chel
30	130.2	62.6	223	22	AAS05209	Mycobacterium scro
31	80	38.5	3495	22	AAH65512	C glutamicum codin
32	80	38.5	349980	22	AAH68525	C glutamicum codin
33	72.4	34.8	27426	23	AAH68525	Propionibacterium
34	43.2	20.8	4956	23	ABL09357	Drosophila melanog
35	43.2	20.8	35633	23	ABL09356	S. tendae nikkomyc
36	40.6	19.5	2918	21	AAZ43922	Streptomyces tende
37	40.6	19.5	2919	21	AAZ44486	Recombinant aquir
38	40.2	19.3	1242	18	AAT78176	M. capsulatus gene
39	40	19.2	2826	24	ABQ90502	M. capsulatus gene
40	40	19.2	49999	24	ABQ90984	50K-cellulase gene
41	39.4	18.9	1894	18	AA766542	Maize glutathione-
42	39	18.8	1100	21	AA776277	DNA sequence encod
43	39	18.8	1167	10	AAAN91083	Maize nitrite redu
44	39	18.8	1844	12	AAQ12000	P. putida KT2440-a
45	38.6	18.6	2142	22	AAF61045	

ALIGNMENTS

RESULT 1
AAS05208
ID AAS05208 standard; DNA; 208 BP.

XX AAS05208;

AC AAS05208;

DT 07-SEP-2001 (first entry)

XX Mycobacterium kansasii rpoB gene fragment.

DE Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;

XX PCR-restriction fragment length polymorphism analysis; ds.

OS Mycobacterium kansasii.

PN WO200131061-A1.

XX 03-MAY-2001.

PF 27-OCT-2000; 2000WO-KR01223.

XX 27-OCT-1999; 99KR-0046795.

PR (ERUM-) ERUME BIOTECH CO LTD.

XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;

XX WPI; 2001-300520/31.

XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -


```
XX Claim 1; Page 42; 50pp; English.
XX
CC The present sequence for Mycobacterium kansasii rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 51 A; 65 C; 65 G; 27 T; 0 other;
Query Match 100.0%; Score 208; DB 22; Length 208;
Best Local Similarity 100.0%; Pred. No. 6.6e-41;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCAAGGAGAAGCGCTACGACCTGCGCCGCTGTCGCGCGGATACAAAGTCAACAAGACTGG 60
DB 1 TCAAGGAGAAGCGCTACGACCTGCGCCGCTGTCGCGCGGATACAAAGTCAACAAGACTGG 60
QY 61 GCCTGAACACCAATCATCCGATCACCACGACGACGCTGACCGAAGAGACGTCGTCGCCA 120
DB 61 GCCTGAACACCAATCATCCGATCACCACGACGACGCTGACCGAAGAGACGTCGTCGCCA 120
QY 121 CCATCGAGTATCTGGTCCGCTGCACGAGGCGCCAGCCAGATGACCGTCCCGGGCGGG 180
DB 121 CCATCGAGTATCTGGTCCGCTGCACGAGGCGCCAGCCAGATGACCGTCCCGGGCGGG 180
QY 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208
DB 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208
RESULT 2
AAS05217
ID AAS05217 standard; DNA; 208 BP.
AC AAS05217;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium celatum rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium celatum.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee He, Park YK, Bal G, Kim S, Cho S, Kim Y, Park HJ;
```

```
XX WPI; 2001-300520/31.
XX
PT New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
PS Claim 1; Page 45; 50pp; English.
XX
CC The present sequence for Mycobacterium celatum rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 48 A; 71 C; 64 G; 25 T; 0 other;
Query Match 83.8%; Score 174.4; DB 22; Length 208;
Best Local Similarity 89.9%; Pred. No. 7.2e-33;
Matches 187; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 1 TCAAGGAGAAGCGCTACGACCTGCGCCGCTGTCGCGCGGATACAAAGTCAACAAGACTGG 60
DB 1 TCAAGGAGAAGCGCTACGACCTGCGCCGCTGTCGCGCGGATACAAAGTCAACAAGACTGG 60
QY 61 GCCTGAACACCAATCATCCGATCACCACGACGACGCTGACCGAAGAGACGTCGTCGCCA 120
DB 61 GCCTGAACACCGCGTCCCGCGATCAGCAGCAGCAGCTCTCACCAGAGAGGACGTCGTCGCCA 120
QY 121 CCATCGAGTATCTGGTCCGCTGTCGCGCGGATGACCGTCCCGGGCGGG 180
DB 121 CCATCGAGTATCTGGTCCGCTGTCGCGCGGATGACCGTCCCGGGCGGG 180
QY 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208
DB 181 TCGAGGTGCCGGTGGAAACCGACGACAT 208
RESULT 3
AAS05215
ID AAS05215 standard; DNA; 208 BP.
XX
AC AAS05215;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium avium rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium avium.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
```

```
XX 27-OCT-1999; 99KR-0046795.
XX (ERUM-) ERUME BIOTECH CO LTD.
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI: 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism -
XX Claim 1; Page 44; 50pp; English.
XX
XX The present sequence for Mycobacterium avium rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 44 A; 69 C; 69 G; 26 T; 0 other;
XX
XX Query Match 82.3%; Score 171.2; DB 22; Length 208;
XX Best Local Similarity 88.9%; Pred. No. 4.2e-32;
XX Matches 185; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
XX
XX 1 TCAAGGAGGAGCGCTACGACCTGGCGCGGTGCGCGGATACAAAGTCAACGAAGCTGG 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 1 TCAAGGAGGAGCGCTACGACCTGGCGCGGTGCGCGGATACAAAGTCAACGAAGCTGG 60
XX
XX 61 GCCTGAACACCAATCATCCGATCACACGACGACGCTGACCGAAGAGACGCTCGTCGCCA 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 GCCTGCACGCCGGTCAGCCGATCACCAGCTCGACGCTGACCGAGGAGACGCTCGTCGCCA 120
XX
XX 121 CCATCGAGTATCTGTCGCTGACGAGGCGGACGACGATGACCGTCCCGGGCGGG 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 121 CCATCGAGTATCTGTCGCTGACGAGGCTGACGCCACGATGACCGTCCCGGGCGGG 180
XX
XX 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 181 TCGAGGTGCGCGTGGAGACCGACGACAT 208
XX
XX RESULT 4
XX AAS05203
XX ID AAS05203 standard; DNA: 208 BP.
XX
XX AC AAS05203;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium gordonae type III rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX PCR-restriction fragment length polymorphism analysis; ds.
XX
XX Mycobacterium gordonae type III.
```

```
XX WO200131061-A1.
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
XX WPI: 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
XX diagnosis and identification of many mycobacterial species by
XX restriction fragment length polymorphism -
XX Claim 1; Page 41; 50pp; English.
XX
XX The present sequence for Mycobacterium gordonae type III rpoB gene
XX fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX various Mycobacterial species. These rpoB gene fragments can be used
XX in the diagnosis and identification of Mycobacterium species using a
XX novel PCR-restriction fragment length polymorphism analysis (PRA)
XX method. The method comprises obtaining a restriction fragment length
XX polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX amplifying and digesting the DNA fragment from the microorganism to
XX be identified and comparing the RFLP patterns from the known rpoB gene
XX fragments with the unidentified fragment. The rpoB gene fragments
XX are useful to identify a wide range of Mycobacterium species, e.g. for
XX diagnosis or to obtain epidemiological and pathogenesis information for
XX selection of appropriate therapies, including M. tuberculosis, M. leprae
XX and non-tuberculous mycobacteria (NTM) encountered in subjects infected
XX with human immunodeficiency virus (HIV). Analysis of the rpoB gene
XX fragments is rapid, precise, simple and cost effective (only 1 PCR
XX required), and can differentiate between many species in a single
XX experiment, including those difficult to distinguish by usual biochemical
XX tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
XX detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 46 A; 69 C; 63 G; 30 T; 0 other;
XX
XX Query Match 81.5%; Score 169.6; DB 22; Length 208;
XX Best Local Similarity 88.5%; Pred. No. 1e-31;
XX Matches 184; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
XX
XX 1 TCAAGGAGGAGCGCTAGGACCTGGCGCGGTGTCGGCGGATACAAAGTCAACGAAGCTGG 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 1 TCAAGGAGGAGCGCTAGGACCTGGCGCGGTGTCGGCGGATACAAAGTCAACGAAGCTGG 60
XX
XX 61 GCCTGAACACCAATCATCCGATCACACGACGACGCTGACCGAAGAGAGCTGTCGCCCA 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 GCCTGCACGCTCGGGGATCCGATCACCAGCTTCACGCTGACCGAAGAGAGCTGTCGCCCA 120
XX
XX 121 CCATCGAGTATCTGTCGCTGACGAGGCGGACGCCACGATGACCGTCCCGGGCGGG 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 121 CCATCGAGTATCTGTCGCTGACGAGGCTGACGACGATGACCGTCCCGGGCGGG 180
XX
XX 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 181 CCGAGGTTCGGGTGGAGACCGACGACAT 208
XX
XX RESULT 5
XX AAS05211
XX ID AAS05211 standard; DNA: 208 BP.
XX
XX AC AAS05211;
XX
XX 07-SEP-2001 (first entry)
XX
```

DE Mycobacterium marinum rpoB gene fragment.
 XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
 KW PCR-restriction fragment length polymorphism analysis; ds.
 XX Mycobacterium marinum.
 OS
 XX WO200131061-A1.
 XX 03-MAY-2001.
 XX 27-OCT-2000; 2000WO-KR01223.
 XX 27-OCT-1999; 99KR-0046795.
 XX (ERUM-) ERUME BIOTECH CO LTD.
 XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
 XX WPI: 2001-300520/31.
 XX
 XX New DNA fragments from the rpoB gene of mycobacteria, useful for
 PT diagnosis and identification of many mycobacterial species by
 PT restriction fragment length polymorphism -
 XX
 XX Claim 1; Page 43; 50pp; English.
 XX
 XX The present sequence for Mycobacterium marinum rpoB gene
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
 CC various Mycobacterial species. These rpoB gene fragments can be used
 CC in the diagnosis and identification of Mycobacterium species using a
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)
 CC method. The method comprises obtaining a restriction fragment length
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
 CC amplifying and digesting the DNA fragment from the microorganism to
 CC be identified and comparing the RFLP patterns from the known rpoB gene
 CC fragments with the unidentified fragment. The rpoB gene fragments
 CC are useful to identify a wide range of Mycobacterium species, e.g. for
 CC diagnosis or to obtain epidemiological and pathogenesis information for
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR
 CC required), and can differentiate between many species in a single
 CC experiment, including those difficult to distinguish by usual biochemical
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
 CC detecting specific Mycobacterial species.
 XX
 SQ Sequence 208 BP; 45 A; 69 C; 68 G; 26 T; 0 other;
 Query Match 81.5%; Score 169.6; DB 22; Length 208;
 Best Local Similarity 88.5%; Pred. No. 1e-31; Mismatches 0; Gaps 0;
 Matches 184; Conservative 0; Indels 24; Indels 0; Gaps 0;
 QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGCGCGGATACAAAGTCAACAAGAAGCTGG 60
 Db 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGCGCGGATACAAAGTCAACAAGAAGCTGG 60
 QY 61 GCCTGAACACCAATATCTCGATACACACGACGCTGACCGAAGAAGACGTCTGTCGCCA 120
 Db 61 GCCTGAACCGCGGCCACGACCATCACAGCTCGACGCTGACCGAAGAAGACGTCTGTCGCCA 120
 QY 121 CCATCCAGTATCTGGTCCGCTGCAGAGGCGCCAGCCAGATGACCGTCCGCGGCGGG 180
 Db 121 CCATCGAATACCTGGTCCGCTTGCACGAGGGCCAGACCCGATGACCGTCCGCGGCGGTG 180
 QY 181 TCGAGGTGCGGCTGGAACCCGACGACAT 208
 Db 181 TCGAGGTGCGGCTCGAGACCGACGACAT 208
 RESULT 6
 AAS05201

ID AAS05201 standard; DNA; 208 BP.
 XX AAS05201;
 XX 07-SEP-2001 (first entry)
 XX Mycobacterium gordonae type I rpoB gene fragment.
 DE Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
 KW PCR-restriction fragment length polymorphism analysis; ds.
 XX Mycobacterium gordonae type I.
 XX WO200131061-A1.
 XX 03-MAY-2001.
 XX 27-OCT-2000; 2000WO-KR01223.
 XX 27-OCT-1999; 99KR-0046795.
 XX (ERUM-) ERUME BIOTECH CO LTD.
 XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
 XX WPI: 2001-300520/31.
 XX
 XX New DNA fragments from the rpoB gene of mycobacteria, useful for
 PT diagnosis and identification of many mycobacterial species by
 PT restriction fragment length polymorphism -
 XX
 XX Claim 1; Page 40; 50pp; English.
 XX
 XX The present sequence for Mycobacterium gordonae type I rpoB gene
 CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
 CC various Mycobacterial species. These rpoB gene fragments can be used
 CC in the diagnosis and identification of Mycobacterium species using a
 CC novel PCR-restriction fragment length polymorphism analysis (PRA)
 CC method. The method comprises obtaining a restriction fragment length
 CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
 CC amplifying and digesting the DNA fragment from the microorganism to
 CC be identified and comparing the RFLP patterns from the known rpoB gene
 CC fragments with the unidentified fragment. The rpoB gene fragments
 CC are useful to identify a wide range of Mycobacterium species, e.g. for
 CC diagnosis or to obtain epidemiological and pathogenesis information for
 CC selection of appropriate therapies, including M. tuberculosis, M. leprae
 CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
 CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
 CC fragments is rapid, precise, simple and cost effective (only 1 PCR
 CC required), and can differentiate between many species in a single
 CC experiment, including those difficult to distinguish by usual biochemical
 CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
 CC detecting specific Mycobacterial species.
 XX
 SQ Sequence 208 BP; 46 A; 71 C; 65 G; 26 T; 0 other;
 Query Match 80.8%; Score 168; DB 22; Length 208;
 Best Local Similarity 88.0%; Pred. No. 2.5e-31;
 Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
 QY 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGGCCGATACAAAGTCAACAAGAAGCTGG 60
 Db 1 TCAAGGAGAGCGCTACGACCTGGCCGCTGTCGGCCGATACAAAGTCAACAAGAAGCTGG 60
 QY 61 GCCTGAACACCAATCATCCGATCACCGACGACGCTGACCGAAGAAGACGTCTGTCGCCA 120
 Db 61 GCCTGACGCTCGCGGATCCGATCACCGACTCCACGCTGACCGAAGAAGACGTCTGTCGCCA 120
 QY 121 CCATCGAGTATCTGGTCCGCTGCACGAGGCGCCAGCGATGACCGTCCGCGGCGGG 180
 Db 121 CCATCGAGTATCTGGTCCGCTGCACGAGGCGCCAGCGATGACCGTCCGCGGCGGCA 180
 QY 181 TCGAGGTGCGGCTGGAACCCGACGACAT 208

```
Db 181 CCGAGTGGCGGTGAGACCGGACGACAT 208
|||||
RESULT 7
AAS05219
ID AAS05219 standard; DNA: 208 BP.
XX
AC AAS05219;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium fortuitum rpoB gene fragment.
XX
PN WO200131061-A1.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
XX WPI; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
XX Claim 1: Page 46; 50pp; English.
XX
XX The present sequence for Mycobacterium fortuitum rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 43 A; 69 C; 69 G; 27 T; 0 other;
XX
XX Query Match 80.8%; Score 168; DB 22; Length 208;
XX Best Local Similarity 88.0%; Pred. No. 2.5e-31;
XX Matches 183; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGGTAGGACCTGGCGGTGTCGCGCGATACAGGTCAACAAGAGCTGG 60
|||||
Db 1 TCAAGGAGAGCGGTAGGACCTGGCGGTGTCGCGCGATACAGGTCAACAAGAGCTGG 60
|||||
Qy 61 GCCTGACACCAATCATCCGATCACCACGACGAGCTGACCGAAGAGAGCTGTCGCCCA 120
|||||
Db 61 GCCTGACACCGCCGCGATCAGCTGTCGACTCTGACCCGAGGAGAGCTGTCGCCCA 120
|||||
```

```
Qy 121 CCATCGAGTATCTGTCGCTGCACGAGGCCAGGCCAGCATGACCGTCCCGGGGGG 180
|||||
Db 121 CCATCGAGTACCTGGTGGCTGCACGAGGCCAGGCCAGCATGACCGTCCCGGGGGG 180
|||||
Qy 181 TCGAGGTGCGCGGTGGAACCGACGACAT 208
|||||
Db 181 TCGAGGTCCCGGTGCGAGGTGACGACAT 208
|||||

RESULT 8
AAS05205
ID AAS05205 standard; DNA: 208 BP.
XX
XX AAS05205;
XX
XX 07-SEP-2001 (first entry)
XX
XX Mycobacterium tuberculosis rpoB gene fragment.
XX
XX Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
KW
XX
XX Mycobacterium tuberculosis.
XX
XX OS WO200131061-A1.
XX
XX PN 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-KR01223.
XX
XX 27-OCT-1999; 99KR-0046795.
XX
XX (ERUM-) ERUME BIOTECH CO LTD.
XX
XX Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX
XX WPI; 2001-300520/31.
XX
XX New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
XX Disclosure; Page 41; 50pp; English.
XX
XX The present sequence for Mycobacterium tuberculosis rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
XX Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
XX
XX Query Match 80.0%; Score 166.4; DB 22; Length 208;
XX Best Local Similarity 87.5%; Pred. No. 5.9e-31;
XX Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGGTAGGACCTGGCGGTGTCGCGCGATACAGGTCAACAAGAGCTGG 60
```

Db 1 TCAAGGAGAGCGTACGACCTGGCCCGTCGGTATAGGTCAACGAAGCTCG 60
Qy 61 GCCTGAACACCAATCATCCGATCACCAGCAGCGCTGACCGAAGAACGACGTCGTCGCCA 120
Db 61 GCCTGCATGTCGGCGAGCCCATCAGCTGTCGACGCTGACCGAAGAACGACGTCGTCGCCA 120
Qy 121 CCATCGAGTATCTGTCGGCTGTCAGAGGCCAGGCCACGATGACCGTCGCGGGGGG 180
Db 121 CCATCGAATATCTGTCGGCTGTCAGAGGGTTCAGACACGATGACCGTTCGCGGGCGG 180
Qy 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208
Db 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208
RESULT 9
AAS05206
ID AAS05206 standard; DNA; 208 BP.
XX
AC AAS05206;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium terrae rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium terrae.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI; 2001-300520/31.
XX
PS New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
PS Claim 1; Page 42; 50pp; English.
XX
CC The present sequence for Mycobacterium terrae rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.
XX
SQ Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;

Query Match 80.0%; Score 166.4; DB 22; Length 208;
Best Local Similarity 87.5%; Pred. No. 5.9e-31;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 1 TCRAGGAGAGCGCTACGACCTGGCCCGTCGGCGGATACAGGTCAACAAGACTCG 60
Db 1 TCRAGGAGAGCGCTACGACCTGGCCCGTCGGTATAGGTCAACAAGACTCG 60
Qy 61 GCCTGAACACCAATCATCCGATCACCAGCAGCGCTGACCGAAGAACGACGTCGTCGCCA 120
Db 61 GCCTGCATGTCGGCGAGCCCATCAGCTGTCGACGCTGACCGAAGAACGACGTCGTCGCCA 120
Qy 121 CCATCGAGTATCTGTCGGCTGTCAGAGGCCAGGCCACGATGACCGTCGCGGGGGG 180
Db 121 CCATCGAATATCTGTCGGCTGTCAGAGGGTTCAGACACGATGACCGTTCGCGGGCGG 180
Qy 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208
Db 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208
RESULT 10
AAS05210
ID AAS05210 standard; DNA; 208 BP.
XX
AC AAS05210;
XX
DT 07-SEP-2001 (first entry)
XX
DE Mycobacterium ulcerans rpoB gene fragment.
XX
KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
KW PCR-restriction fragment length polymorphism analysis; ds.
XX
OS Mycobacterium ulcerans.
XX
PN WO200131061-A1.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-KR01223.
XX
PR 27-OCT-1999; 99KR-0046795.
XX
PA (ERUM-) ERUME BIOTECH CO LTD.
XX
PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI; 2001-300520/31.
XX
PS New DNA fragments from the rpoB gene of mycobacteria, useful for
PT diagnosis and identification of many mycobacterial species by
PT restriction fragment length polymorphism -
XX
PS Claim 1; Page 43; 50pp; English.
XX
CC The present sequence for Mycobacterium ulcerans rpoB gene
CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
CC various Mycobacterial species. These rpoB gene fragments can be used
CC in the diagnosis and identification of Mycobacterium species using a
CC novel PCR-restriction fragment length polymorphism analysis (PRA)
CC method. The method comprises obtaining a restriction fragment length
CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
CC amplifying and digesting the DNA fragment from the microorganism to
CC be identified and comparing the RFLP patterns from the known rpoB gene
CC fragments with the unidentified fragment. The rpoB gene fragments
CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC fragments is rapid, precise, simple and cost effective (only 1 PCR

CC required), and can differentiate between many species in a single
CC experiment, including those difficult to distinguish by usual biochemical
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

XX SQ Sequence 208 BP; 45 A; 68 C; 67 G; 28 T; 0 other;
Query Match 80.0%; Score 166.4; DB 22; Length 208;
Best Local Similarity 87.5%; Pred. No. 5.9e-31;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGACCTGCGCCGCTGTCGCGCGATACAAAGTCAACAAGAGCTGG 60
Db 1 TCAAGGAGAGCGCTACGACCTGCGCTGCGTGGTACAAAGTCAACAAGAGCTCG 60
Qy 61 GCCTGAACACCATCATCCGATCACCAGCAGCAGCTGACCGAAGAGAGCTGTCGCCA 120
Db 61 GCCTGAACCGCGCCAGCCATCACCAGCTCGACGCTGACCGAGAGAGAGCTGTCGCCA 120
Qy 121 CCATCGAGTATCTGCTCGGCTGCACGAGGCCAGGCCAGCATGACCGTCCCGGGGGG 180
Db 121 CCATCGAATACCTGCTCGGCTGTCACGAGGCCAGGCCAGCGGATGACCGTCCCGGGGGT 180
Qy 181 TCGAGGTCCCGTGGAAACGACGACAT 208
Db 181 TCGAGGTCCCGTGGAGACCGAGACAT 208

RESULT 11
AAS05216
ID AAS05216 standard; DNA; 208 BP.
XX AC AAS05216;
XX DT 07-SEP-2001 (first entry)
XX DE Mycobacterium bovis rpoB gene fragment.

XX KW Non-tuberculous mycobacteria; rpoB gene fragment; NTM; HIV; PRA; RFLP;
XX KW PCR-restriction fragment length polymorphism analysis; ds.
XX OS Mycobacterium bovis.

XX PN WO200131061-A1.
XX PD 03-MAY-2001.
XX PF 27-OCT-2000; 2000WO-KR01223.
XX PR 27-OCT-1999; 99KR-0046795.
XX PA (ERUM-) ERUME BIOTECH CO LTD.
XX PI Lee H, Park YK, Bai G, Kim S, Cho S, Kim Y, Park HJ;
XX WPI; 2001-300520/31.

XX PT New DNA fragments from the rpoB gene of mycobacteria, useful for
XX PT diagnosis and identification of many mycobacterial species by
XX PT restriction fragment length polymorphism -
XX PS Claim 1; Page 45; 50pp; English.
XX CC The present sequence for Mycobacterium bovis rpoB gene
XX CC fragment is 1 of 24 rpoB gene fragments (AAS05201-AAS05224) from
XX CC various Mycobacterial species. These rpoB gene fragments can be used
XX CC in the diagnosis and identification of Mycobacterium species using a
XX CC novel PCR-restriction fragment length polymorphism analysis (PRA)
XX CC method. The method comprises obtaining a restriction fragment length
XX CC polymorphism (RFLP) pattern of the 24 rpoB gene fragments; isolating,
XX CC amplifying and digesting the DNA fragment from the microorganism to
XX CC be identified and comparing the RFLP patterns from the known rpoB gene
XX CC fragments with the unidentified fragment. The rpoB gene fragments

CC are useful to identify a wide range of Mycobacterium species, e.g. for
CC diagnosis or to obtain epidemiological and pathogenesis information for
CC selection of appropriate therapies, including M. tuberculosis, M. leprae
CC and non-tuberculous mycobacteria (NTM) encountered in subjects infected
CC with human immunodeficiency virus (HIV). Analysis of the rpoB gene
CC fragments is rapid, precise, simple and cost effective (only 1 PCR
CC experiment), and can differentiate between many species in a single
CC tests. Also described are oligonucleotide probes (AAS05227-AAS05242) for
CC detecting specific Mycobacterial species.

XX SQ Sequence 208 BP; 46 A; 62 C; 67 G; 33 T; 0 other;
Query Match 80.0%; Score 166.4; DB 22; Length 208;
Best Local Similarity 87.5%; Pred. No. 5.9e-31;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 TCAAGGAGAGCGCTACGACCTGCGCCGCTGTCGCGCGATACAAAGTCAACAAGAGCTGG 60
Db 1 TCAAGGAGAGCGCTACGACCTGCGCCGCTGCGTGGTACAAAGTCAACAAGAGCTCG 60
Qy 61 GCCTGAACACCATCATCCGATCACCAGCAGCAGCTGACCGAAGAGAGCTGTCGCCA 120
Db 61 GCCTGAATGTCGCGAGCCCATCAGCTGCTGACCGAAGAGAGCTGTCGCCA 120
Qy 121 CCATCGAGTATCTGCTCGGCTGCACGAGGCCAGGCCAGCATGACCGTCCCGGGGGG 180
Db 121 CCATCGAATATCTGCTCGGCTGTCACGAGGGTCAACACCGACGATGACCGTCCCGGGGGG 180
Qy 181 TCGAGGTCCCGTGGAAACGACGACAT 208
Db 181 TCGAGGTCCCGTGGAAACCGACGACAT 208

RESULT 12
AAT09676
ID AAT09676 standard; DNA; 970 BP.
XX AC AAT09676;
XX DT 15-OCT-1996 (first entry)
XX DE Mycobacterium tuberculosis rpoB gene DNA sequence.

XX KW Tuberculosis; disease diagnosis; oligonucleotide; DNA primer; PCR;
XX KW polymerase chain reaction; DNA amplification; rpoB locus; TB; ss.

XX OS Mycobacterium tuberculosis.
XX FH Key Location/Qualifiers
XX FT primer_bind 10..27
XX FT /tag= a
XX FT /note= "primer FENLFF"
XX FT primer_bind 226..243
XX FT /tag= b
XX FT /note= "primer DDIDLH"
XX FT primer_bind 226..240
XX FT /tag= c
XX FT /note= "primer DDIDLH"
XX FT primer_bind 338..364
XX FT /tag= d
XX FT /note= "primer rpo95"
XX FT primer_bind 348..373
XX FT /tag= e
XX FT /note= "primer rpo105"
XX FT primer_bind 354..373
XX FT /tag= f
XX FT /note= "primer KY290"
XX FT misc_feature 372..373
XX FT /tag= g
XX FT /note= "M. tuberculosis signature nucleotide"
XX FT misc_feature 433..434
XX FT /tag= h


```
SQ Sequence 3519 BP; 675 A; 1081 C; 1183 G; 580 T; 0 other;
Query Match      80.0%; Score 166.4; DB 22; Length 3519;
Best Local Similarity 87.5%; Pred. No. 7.4e-31;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 1 TCAAGGAGAAGCGCTAGGACCTGGCCCGTGTGCGCGGATACAAAGGTCAACAAGAGCTGG 60
DB 884 TCAAGGAGAAGCGCTAGGACCTGGCCCGTGTGCGGTCTAAGGTCAACAAGAGCTCG 943
OY 61 GCCTGAACACCAATCATCCGATCACCACGAGGAGCTGACCGAAGAGAGCGTGTGCGCCA 120
DB 944 GGCTGCATGTGCGGAGCCCATCAGCTGCTGACGCTGACCGAAGAGAGCGTGTGCGCCA 1003
OY 121 CCATCGAGTATCTGCTCGGCTGCACGAGGCCAGGCGCACGATGACCGTCCCGGGGGG 180
DB 1004 CCATCGAATATCTGCTCGGCTGCACGAGGGTCAGACCAACGATGACCGTTCGCGGGGGG 1063
OY 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208
DB 1064 TCGAGGTGCGCGTGGAAACCGACGACAT 1091

RESULT 14
ID AAH02079 standard; DNA: 3534 BP.
XX AC AAH02079;
XX DT 24-JUL-2001 (first entry)
XX DE Mycobacterium tuberculosis nucleotide sequence SEQ ID NO:2072.
XX KW Species specific; genus specific; family specific; probe: detection;
XX KW identification; algal; archaeal; bacterial; fungal; parasitica;
XX KW microorganism; diagnosis; translation elongation factor Tu; toxin;
XX KW translation elongation factor G; RecA recombinase; resistance;
XX KW catalytic subunit of proton-translocating ATPase; antimicrobial;
XX KW vaccine; primer; ds.
XX OS Mycobacterium tuberculosis.
XX PN WO200123604-A2.
XX PD 05-APR-2001.
XX PF 28-SEP-2000; 2000WO-CA01150.
XX PR 28-SEP-1999; 99CA-2283458.
XX PR 19-MAY-2000; 2000CA-2307010.
XX PA (INFE-) INFECTIO DIAGNOSTIC (IDI) INC.
XX PI Bergeron MG, Boissinot M, Huletsky A, Menard C, Ouellette M;
XX PI Picard FJ, Roy PH;
XX DR WPI: 2001-245006/25.
XX PT Nucleic acid sequences are used to generate universal probes and
XX PT primers which can be used to identify and detect the presence of algal,
XX PT archaeal, bacterial, fungal and parasitica species in a test sample -
XX PS Disclosure; Page 1478-1479; 1580pp; English.
XX CC The present invention describes a method for generating a repertory of
XX CC nucleic acids of tuf, fus, atpD and/or recA genes from which probes
XX CC and/or primers are derived. The method comprises amplifying the nucleic
XX CC acids of determined algal, archaeal, bacterial, fungal and parasitica
XX CC species with a combination of defined primer pairs. The method can be
XX CC used for producing probes and/or primers for detecting one or more
XX CC related microorganisms e.g. algae, archaea, bacteria, fungi and
XX CC parasites, for universal detection and for specific and ubiquitous
XX CC detection and identification of an algal, archaeal, bacterial, fungal
```

```
CC and parasitica species, genus, family and group. A nucleic acid (1)
CC obtained using the method of the invention can be used for the universal
CC detection of any bacterium, fungus or parasite in a sample and for the
CC detection of at least one antimicrobial agent resistance gene or at
CC least one toxin gene. hexA nucleic acids are used for the specific and
CC ubiquitous detection and for identification of Streptococcus pneumoniae.
CC (1) can be used to design a therapeutic agent which is effective against
CC microorganisms. Microbial species or genus or family or phylum or group
CC which can be detected include Abiotrophia adiacens, Bordetella sp.,
CC Corynebacterium sp., Enterobacteriaceae group, Escherichia coli,
CC Mycobacteriaceae family, Pseudomonads group, Streptococcus sp.,
CC Neisseria gonorrhoeae and Staphylococcus sp.. Using DNA based tests
CC provides faster results than substrate specificity tests as results can
CC be determined in an hour and improved accuracy is also achieved.
CC AAH00010 to AAH002304 represent nucleotide sequences and primers/probes
XX CC which are given in the exemplification of the present invention.
XX SQ Sequence 3534 BP; 679 A; 1081 C; 1188 G; 586 T; 0 other;
Query Match      80.0%; Score 166.4; DB 22; Length 3534;
Best Local Similarity 87.5%; Pred. No. 7.4e-31;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 1 TCAAGGAGAAGCGCTAGGACCTGGCCCGTGTGCGCGGATACAAAGGTCAACAAGAGCTGG 60
DB 902 TCAAGGAGAAGCGCTAGGACCTGGCCCGTGTGCGGTCTAAGGTCAACAAGAGCTCG 961
OY 61 GCCTGAACACCAATCATCCGATCACCACGAGGAGCTGACCGAAGAGAGCGTGTGCGCCA 120
DB 962 GGCTGCATGTGCGGAGCCCATCAGCTGCTGACGCTGACCGAAGAGAGCGTGTGCGCCA 1021
OY 121 CCATCGAGTATCTGCTCGGCTGCACGAGGCCAGGCGCACGATGACCGTCCCGGGGGG 180
DB 1022 CCATCGAATATCTGCTCGGCTTGCACGAGGGTCAGACCAACGATGACCGTTCGCGGGG 1081
OY 181 TCGAGGTGCGCGTGGAAACCGACGACAT 208
DB 1082 TCGAGGTGCGCGTGGAAACCGACGACAT 1109

RESULT 15
ID AAA74651 standard; DNA: 3853 BP.
XX AC AAA74651;
XX DT 06-DEC-2000 (first entry)
XX DE Mycobacterium tuberculosis rpoB gene.
XX KW Mycobacterium tuberculosis; rpoB; RNA polymerase beta subunit;
XX KW rifampin resistance; mutation detection; ds.
XX OS Mycobacterium tuberculosis.
XX PN WO200043546-A2.
XX PD 27-JUL-2000.
XX PF 20-DEC-1999; 99WO-US30377.
XX PR 19-JAN-1999; 99US-0233996.
XX PA (DADE-) DADE BEHRING INC.
XX PI Liu YP, Kurn N;
XX DR WPI: 2000-524243/47.
XX PT Method for detecting drug resistance in a strain of an organism,
XX PT particularly for detecting rifampin resistance in Mycobacterium
XX PT tuberculosis -
```


PS Example 1; Fig 4; 86pp; English.

xx
CC The present sequence is the rpoB gene from Mycobacterium tuberculosis.
CC Rifampin resistance is largely associated with point mutations
CC localised in a small core region of 81 base pairs in the rpoB gene, which
CC encodes the RNA polymerase beta subunit. To detect a mutation, a complex
CC is formed comprising a first sequence representing the predetermined
CC region of the gene of the organism and a second sequence representing the
CC corresponding region of the gene of the wild type organism in double
CC stranded form. Each member of at least one pair of non-complementary
CC strands within the complex has a label. The association of the labels in
CC the complex is related to the presence of the mutation. The presence of
CC the mutation is related to the drug resistance of the strain.
xx

SQ Sequence 3853 BP; 723 A; 1173 C; 1293 G; 564 T; 0 other;

Query Match 80.0%; Score 166.4; DB 21; Length 3853;
Best Local Similarity 87.5%; Pred. No. 7.5e-31;
Matches 182; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 1 TCAAGGAGACGCTACGACCTGGCCGCTGTCGGCCGATACAGGTCAACAAGAAGCTGG 60
Db 1477 TCAAGGAGACGCTACGACCTGGCCGCTGTCGGCTCGTATAGGTCAACAAGAAGCTCG 1536
Qy 61 GCCTGAACACCAATCATCCGATCACACGACGCTGACCGAAGAGAGAGCTGTCGCCA 120
Db 1537 GGCTGCATGTCGGCGAGCCCATACGTCGTGACGCTGACCGAAGAGAGAGCTGTCGCCA 1596
Qy 121 CCATCGAGTATCTGGTCGGCTGCACGAGGCCAGGCCAGCATGACCGTCCGGGGGGG 180
Db 1597 CCATCGAATATCTGGTCGGCTTGCACGAGGGTCAGACGATGACCGTTCCGGGGGGG 1656
Qy 181 TCGAGTGCCGGTGGAAACCGACACAT 208
Db 1657 TCGAGTGCCGGTGGAAACCGACACAT 1684

Search completed: November 12, 2002, 16:50:09
Job time : 141.222 secs